IRE Transactions

on Reliability and Quality Control

PGRQC-12

NOVEMBER, 1957

TABLE OF CONTENTS

A Redundancy Analog...A. C. Block 1

The Reliability Qualification of Electronic Equipment..............D. W. Pertschuk 8

Passive Components for Submarine Telephone Cable Repeaters..M. C. Wooley 14

Statistical Aspects of Reliability in Systems Development..........J. S. Youtcheff 25

A Sequential Test for Comparing Component Reliabilities..........C. F. Stevens 37

PUBLISHED BY THE

PROFESSIONAL GROUP ON RELIABILITY AND QUALITY CONTROL

IRE PROFESSIONAL GROUP ON RELIABILITY AND QUALITY CONTROL

The Professional Group on Reliability and Quality Control is an organization, within the framework of the IRE, of members with principal professional interest in Reliability and Quality Control. All members of the IRE are eligible for membership in the Group and will receive all Group publications upon payment of prescribed assessments.

Annual Assessment: $2.00

IRE TRANSACTIONS®
on Reliability and Quality Control

Published by the Institute of Radio Engineers, Inc., for the Professional Group on Quality Control, 1 East 79th Street, New York 21, New York. Responsibility for the contents rests upon the authors, and not upon the IRE, the Group or its members. Individual copies available for sale to IRE-PGRQC members at $0.90, to IRE members at $1.35 and to nonmembers at $2.70.

A REDUNDANCY ANALOG

Alfred C. Block
Dynamics Research Associates
Ferguson, Missouri

In the literature of reliability the subject of redundancy plays an important role. Many authors point out that redundancy may be the only means of achieving a given reliability when given state-of-the-art components can not do the job.[1,2]

The question of redundancy can be shown to be one of degree. This paper attempts to establish the degree and present principles, developed by analogy to natural systems, to guide the designer in specifying redundancy.

REDUNDANCIES IN PRACTICE

Consider a complex mechanism, such as an airplane, which is required by specification to have a reliability of 90 per cent. If it is only possible to build an airplane whose reliability (in fact) is shown to be 65 per cent, we can increase the reliability by redundancy. Obviously, the consideration of redundancy might lead to the duplication of the entire airplane and the use of more airplanes to accomplish the task. This use of redundancy might be termed gross redundancy and has little application. An example of the nonusefulness of this redundancy is seen if we specify a transport airplane to have a 0.99 chance of carrying 50 passengers from New York to London. This is not the same as two airplanes each having a 0.90 chance of carrying 50 passengers. In this case the redundancy even applies to the passengers. We need 100 passengers to assure the arrival of at least 50 within the specified reliability limit. This is a rather absurd use of redundancy, especially from a passenger's standpoint. Yet when expendable items such as bombs are specified for the pay-load, this is the reasoning that many people apply. It does serve warning that there is something more to be considered than the feeling that redundancy is a universal panacea to reliability problems.

The more sensible approach, indeed the one which most people would choose, is to seek out the offending member causing this low reliability and duplicate it within the airplane. In fact, one might do more and search for the culprit within this offending system and narrow the duplication to a low reliability component. This might take some detective work on the order of the BOAC effort to determine the reason for their Comet crashes, but the net result would be beneficial. This other extreme in the application of redundancy might be termed minute redundancy.

Between these two extremes is the sought-for redundancy. The evaluation of how much and where to apply it must be bound by considerations outside redundancy itself.

REDUNDANCY INVOLVING SWITCHING

Let us consider first one of the most serious objections to the application of redundancy, the necessity of switching. If a system is duplicated and only one is in use at a time, a switch must be added to place the redundant system in

a functional position, when doing this would depend upon some decision-making body (e.g., the system operator). Generally, the second system is switched in only when there is a failure in the first system.

Before any switching should be done, knowledge to make the decision to switch must be had (i.e., resolving the question of failure in the first system before switching takes place). In other words, the failure must be knowledgeable. Hence, for any redundancy requiring switching, some knowledge must be necessary. This knowledge must be obtained from outside the implied redundancy and may be a failure indicator or a use doctrine (e.g., "switch to set II after 10 hours").

As a consequence, any switch used to determine which redundant element should be used implies a set of other instruments such as a failure indicator and a timer, not mutually exclusive, although only one such instrument may be necessary. A failure detector may be a complex device. By using redundancy as a method of increasing reliability, we may have added more complication with its corollary, lower reliability, to our total system.

REDUNDANCY AND MAINTENANCE

Redundancy adds additional problems in maintaining equipment. Gross redundancy always involves maintaining twice as much equipment. Minute redundancy always involves time factors. For instance, we must know when each of the components is ready for replacement. Spares must be on hand for such replacement. For gross redundancy the storage, amount and the use of the spares involves large movements and large amounts of men and materials. The logistics problem of any operational base is intensified. For minute redundancy, the spares must be instantly available, in fixed quantities and of more than ample amounts. Times of use records must be kept so that replacements are made on schedule.

It can be seen that the very things which are trying to be saved are being worsened. An interesting aspect of this peculiarity is the case of redundancy where an automatic switching device or an inherent redundancy is used. In this case the user may not know of a failure in a primary unit. This is no worse than minute redundancy, but the effect is much different. One whole equipment has failed rather than a duplicate component, and the remaining capability is unitized rather than quantitized. Yet at each initial use time, it is expected that start-up is at maximum capability. This implies that after each use period the failure must be found and repaired even when there is no apparent functional failure.

REDUNDANCY AND COST

For gross redundancy, the additional cost is directly proportional to the cost of a single system. It may be expressed by a simple formula:

$$C_{rg} = n\, C_o \tag{1}$$

where C_{rg} = the cost of the entire grossly redundant system
n = the number of duplicated systems
C_o = the cost of the single nonredundant system.

For a minute redundant system the cost may be expressed by:

$$C_{rm} = \sum_{i=1}^{k} (a_i - c_i) \tag{2}$$

where C_{rm} = the cost of the minutely redundant system
a_i = the additional cost of i^{th} component because of the built-in redundancy
c_i = the original cost of the i^{th} component
k = the number of components.

Note that

$$C_o = \sum_{i=1}^{k} c_i \,. \tag{3}$$

For systems which are redundant by duplicating subsystems the cost formula may be written as:

$$C_{rs} = C_o + \sum_{i=1}^{n} S_i + c_s \tag{4}$$

where C_{rs} = the cost of the redundant systems
S_i = the cost of i^{th} system to be duplicated
c_s = the cost of switches and indicators, etc., added.

It does not necessarily follow that any inequalities may be drawn from the above formulas. Several authors have discussed redundancy from the cost standpoint, and in almost every case the question of cost proves to be tied to an operations research type of analysis.[3,4]

The simple formulas above are presented to show statements of how costs would ordinarily be computed and to show the way the new costs are added.

ALTERNATE MODES

A type of redundancy often used is the method of alternate modes. This involves the use of an alternate method of accomplishing the desired function. Often, the secondary method is performed at reduced effectiveness. A case in point is a fire control system, whose primary sighting device is radar. If the radar capability is lost, the operator may switch to optical sighting, although this is of little value at night. Why the optical sight is incorporated rather than duplicating the radar sight is debatable. Perhaps the answer is tied up to the "holding on" of something old. This tendency in design is fully discussed by Henry Dreyfuss, a successful industrial designer.[5] He states, relative to his designs:

> "Almost without exception, our designs include an ingredient we call survival form. We deliberately incorporate into the product some remembered detail that will recall to the users a similar article put to a similar use....Somehow these recollections of the past give us comfort, security, and silent courage. By embodying a familiar pattern in an otherwise wholly new and possibly radical form, we can make the unusual acceptable to many people who would otherwise reject it."

In the case of the fire control system, the optical sight is an ancient survival form. It should be a rule for new designs to free them from hide-bound opinions. Certainly, no redundancy which serves to lower the effectiveness should

be considered merely because there are operators who like the old way best. The answer is simple: train the operators in the new way. A retraining period for the operators which may involve detraining first is sure to pay dividends in terms of usefulness and effectiveness when the equipment is needed. In terms of the job to be performed, the emphasis on new designs should not be held back by the operator's initial training. Unfortunately, the preconceived notions of experienced personnel are very important items in the selection of new equipment. A proper outward orientation must be established at the onset of specification writing period.

THE CASE AGAINST REDUNDANCY

It is a matter of concern to hear the redundancy argument applied categorically. It is not the intent of this paper to deny the use of redundancy. However, the use of redundancy is a mixed blessing and must be used with knowledge and discretion. What looks like a case against redundancy is, in reality, a plea for the judicious use of redundancy. The criteria governing this use are in fact simple and direct.

BIOLOGICAL REDUNDANCY

Many authors have pointed out that nature does not depend upon a series system. The human race is an example of a very highly redundant system. In like manner, one human being is a highly redundant system, and it is of extreme interest to see how this redundancy is applied.

Consider two of the most important organs, the heart and the brain. Neither of these are examples of gross redundancy as compared to the lungs or kidneys. On the other hand they are composed of redundant elements, although not even in the next step down (subsystem). In the brain we find only one cerebrum and only one cerebellum. In the heart we find two auricles and ventricles, but the process of use of each is such that they are used in series. At the next step downward we run into component (minute) redundancy. All the biological cells (muscle fibers or brain tissue) are duplicated in extremum.

Many seeming redundancies in the human body are not redundancies at all (e.g., two eyes). But how and why did nature evolve so that the redundancy of lungs, kidneys, etc., exist? Perhaps we can learn from this and build our own equipment in the same way. Nature has some restorative power. It is limited in the case of large animals to very small portions (i.e., cells). For simple animals, such as worms, restoration of destructed organs is complete.

With electronic equipment, restoration can be made for almost any part from the smallest to the entire system. In many cases function is preserved or curtailed for both during the restorative process. We call this restoration for our inanimate equipment, maintenance, and the ability to make restoration, maintainability.

INITIAL APPLICATION OF REDUNDANCY

Redundancy application should start with the simplest device that can be duplicated. This may be a part of a component, such as the much used example of a twin contact relay, or it may be a small component itself, such as a rivet. What

should be looked for is cell reliability. In the design of equipment, this type of redundancy can be easily applied. Switches, for example, can be used redundantly or have redundancies in them. Certainly, in comparing two individual switches, the one with the redundancies would be assumed to be the better one from a reliability standpoint. If the redundant switch is itself used redundantly, a tremendous improvement in reliability could be obtained. This principal is used necessarily (e.g., in making cables). The correct application of series or parallel redundancy is much easier at the cell level than any "higher" level of building. The question of cost in cell duplication is nominal whereas in black-box duplication it may be disproportionate. This is exemplified by the self-checking reciprocal circuits in the large computing machines such as Harvard's Mark II or Remington Rand's Univac. However, for equal reliabilities the cost of a many-celled system vs. a many black-boxed system would be subject to investigation. In cases where reliability is of paramount importance, cost is subject to future considerations, and the question of first cost vs. continuing costs must be looked into. The continuing cost is a factor no matter how the equipment is built. Therefore, a better means of measurement would be the application of operations research to determine effectiveness vs. cost, and this should be the main parameter upon which decisions are based.

Effectiveness as used here is defined as the product of performance and reliability, where performance is rated on the same scale showing the ratio of actual to planned accomplishment and reliability is as defined in the literature.[1,2]

A REDUNDANCY TEST

One question which arises is the correct (or best) application of redundancy for equal reliability. Assuming that a given reliability could be achieved by either minute or gross redundancy, is there some measure that can be used to determine which is best?

Some answer may be found in the application of information theory to reliability. It has been shown that entropy is a suitable measure to determine choice for a set of equal reliability.[6]

This measure may be applied to redundant systems as well and shows, for a given fixed set, a lower entropy; i.e., best choice, for a refined system rather than a gross system. This stems from the fact that parts containing many cells do not increase the complexity of the system (see Appendix).

VULNERABILITY

Using the human body as an example again, we go back to the question of why one heart but two lungs, why one liver but two kidneys, etc. One inference that may be drawn is that of vulnerability. If the physical function of the human body is examined in terms of a natural surrounding, and the type of accident or failure that can occur is limited only by those which can be inflicted by blows, we can easily see that the gross redundancy is used in vulnerable areas and not as a function of necessity for vital performance. The heart is protected by the sternum, a heavy chest bone armor, and the lungs are correspondingly protected by the ribs. But lung accidents because of rib puncture are fairly common, whereas a heart failure because of bone breakage from body blows is extremely rare. Simi-

larly, for liver and kidneys. The liver is well protected, but the kidneys are not.

The analogy to equipment is easily made. A vulnerable part is one which is subject to environmental failure. Improvement may be made in a system to increase reliability by making the item grossly redundant or by changing its environment. The fix will be dictated by circumstances. Again the point is, a little thought before the action will pay high dividends later. The categorical approach of redundancy application may not be the answer.

CONCLUSION

The proper application of redundancy is as yet in its infancy. We are concerned here with defining the terms and formalizing a concept. We have shown that there is a redundancy spectrum and that the choice of the proper type of redundancy lies with the designer. No categorical approach may be used but a carefully planned program of reliability improvement should be followed starting with the simplest device to achieve cell reliability and ending with the system itself. Intermediate problems are handled as they occur and in many cases the answer is not redundancy but simply changing external effects. Where redundancy proves necessary, a choice for first, and not alternate, methods should be made, and the question of cost should be related to effectiveness.

ACKNOWLEDGEMENT

The author wishes to acknowledge the help, advice and encouragement of D. L. Foley of the Boeing Airplane Company and E. D. Goddess, Management Consultant, Seattle, in the preparation of this paper.

APPENDIX

Using the methods of A. C. Block,[6] we will establish the best choice of redundancy for two similar systems. Let the original (nonredundant) system have a reliability P_1 and be composed of n elements each having a reliability p_i. Then the entropy, H_1, or measure of choice is[6]

$$H_1 = - \sum_{i=1}^{n} p_i \log p_i \tag{1}$$

for the element which contains a built-in redundancy, $p_{jr} > p_i$. If some, say k of them, of the p_i's are replaced by p_{jr}'s then Eq. (1) is replaced by

$$H_2 = - \sum_{i=1}^{k} p_{jr} \log p_{jr} + \sum_{i=k}^{n} p_i \log p_i . \tag{2}$$

But since the individual $p_{jr} > p_i$, $H_2 < H_1$ and P_1 is replaced by P_2 where $P_2 < P_1$. Now let the original system be duplicated in its entirety, so that $P_2 = 1 - (1 - P_1)^2$. Then $H_2 - 2H_1$ since

$$H_2 = - \sum_{i=1}^{2n} p_i \log p_i . \tag{3}$$

In this case $H_2 > H_1$ and the best choice, based upon minimum entropy, is for that system which has built-in redundancies.

BIBLIOGRAPHY

1. Carhart, R. R.: "Survey of the current status of the electronic reliability program." Rand Corporation Report RM-1131. (Aug. 14, 1953).

2. Lindsay, R. H.: "Practical expectations and limitations of the reliability problem." Aeronautical Engineering Review 13:65-71 (Oct. 1954).

3. DeToro, M. J.: "Reliability criterion for constrained systems." Trans. IRE, PGRQC-8, pp. 1-6 (Sept. 1956).

4. Moskowitz, F., and McLean, J. B.: "Some reliability aspects of systems design." Trans. IRE, PGRQC-8, pp. 7-35 (Sept. 1956).

5. Dreyfuss, H.: "Designing for People." New York, Simon and Schuster, 1955, pp. 59-60.

6. Block, A. C.: "An entropic measure of reliability." Transactions, Western Quality Control Conference, ASQC, Los Angeles (Aug. 1956).

THE RELIABILITY QUALIFICATION OF ELECTRONIC EQUIPMENT

David W. Pertschuk
Arma Division
American Bosch Arma Corporation
Garden City, New York

There are not many of us left in the electronics business who can afford to ignore the problem of reliability. And those happy few who are left, certainly are not in the business of designing aeronautical or navigational equipment. Reliability is a large and serious problem and, in some cases, one of life and death. For the electronics industry as a whole, reliability is a limiting and constraining barrier to growth. Our ability to design circuits has too often in the past exceeded our ability to design hardware which will operate without failure for a sufficient length of time. So nearly all of us, to a lesser or greater extent, are concerned with the reliability of the equipment we design and manufacture.

Our concern with reliability, however, does not disguise the fact that it is only one of the characteristics of the equipment. It is not too difficult to achieve high reliability if we are permitted to neglect such factors as performance, weight, size, cost and development time. Since we obviously cannot neglect these factors, the problem becomes one of finding the right compromise between reliability and the other necessary characteristics of the equipment. If this compromise is to be made in some purposeful and logical way, and if, above all, we are to learn from experience, it is clear that we cannot long continue to work in a hit-and-miss fashion. We must sooner or later develop methods for the reliability qualification of our end products.

Reliability evaluation cannot be permitted to depend entirely on customer reaction. It is true, of course, that the customer's opinion of the reliability of our equipment is always the acid test. However, by the time this opinion has been formed and the information has trickled back to the manufacturer, remedial action is, at best, excessively difficult and exorbitantly expensive. It is the purpose of this paper to discuss methods of assessing the reliability of electronic equipment before large-scale delivery to the customer is made.

Suppose we have been given the job of designing and producing a piece of electronics equipment, a "black box," which will become part of the stabilization system of a new high-performance military aircraft. Assume that this black box is a new concept -- nothing quite like it has been made before. We are informed that the reliability of the equipment is important because its failure, while not catastrophic in the sense that the aircraft will crash, will certainly result in the tactical abortion of the flight. We are given detailed and complete performance specifications, some rough idea of the external environment within which our system must live and the length of time the system must operate during a mission. The desired reliability goal is 0.97, and a reliability of less than 0.84 will not be acceptable. This is interpreted to mean that, in the long run, and at the very worst, our equipment must operate adequately in at least 84 out of every 100 flights, and that it is desirable that this rate be as close as possible to 97 out of every 100 flights.

For some of the reasons just discussed, it is of considerable importance to us to demonstrate that our finished product has met the desired level of reliability. The most immediate difficulty we encounter in attempting to arrange such a demonstration is the problem of evaluating "long-run" reliability. We have not been asked to produce our equipment in such a way that out of the first 100 units delivered at least 97 operate properly, nor even that out of the first 1,000 at least 970 operate properly. Instead, we have been told, to put it mathematically, that as the number of units we produce approaches infinity, the proportion of successes should approach 97 per cent. Clearly this point can never be proven exactly with any finite test. As a consequence, both the customer and ourselves must be willing to assume certain risks.

It is entirely possible for the equipment to have the desired reliability and for a finite test to fail to demonstrate this fact. We shall call this the manufacturer's risk. Conversely, it is possible for the equipment to lack the desired reliability and for the test to indicate that the reliability is sufficiently high. We shall call this the customer's risk. In order to reduce these risks to acceptably low levels, it is not only necessary that any demonstration be very carefully planned, but that the desired reliability of a product be given in terms of two figures. The upper one, in our example 0.97, should be sufficiently high for it to be unlikely that a demonstration will reject a product of this high reliability. The lower one, in our example 0.84, should be sufficiently low for it to be unlikely that a demonstration will accept a product of this low reliability. This will explain why any enlightened reliability specification, made either by the customer or by the designer himself, should be given in terms of an upper and a lower limit.

Assume now that the system has been designed, the design functionally evaluated on breadboards, prototype packages built and the final production system is starting to come off the assembly line. Reliability has been given high priority throughout these development and early production stages and, without going into details, we shall suppose that the manufacturer is satisfied that his product will meet its reliability specification. It now becomes necessary for the equipment to receive some form of reliability qualification or evaluation.

Now let us design an ideal reliability qualification for our imaginary black box. How many units should we test, and how many should work properly? If we give any competent statistician our reliability limits and tell him that we have fixed the manufacturer's and customer's risks both at 5 per cent, he will come up with a test plan in the form of an operating characteristic (O-C) curve. Figure 1 shows the curve we shall need. It can be seen that a sample of 46 units is necessary and that the test will be considered successful if at least 43 units pass. The O-C curve plots the probability of the test's being successful against the actual reliability of the equipment. It is clear that our manufacturer's and customer's risk requirements have been satisfied. There is a 95 per cent chance of the test's being successful if the actual equipment reliability is as high as 0.97; there is only a 5 per cent chance if the actual equipment reliability is as low as 0.84. The curve also shows the probability of a successful test for all values of reliability between 0.84 and 0.97.

Having settled the statistical details, we must now consider the actual form of the test to be given each of the units in our sample. Here again, the answer is a fairly obvious one. We should give each of the units an exact simulation of

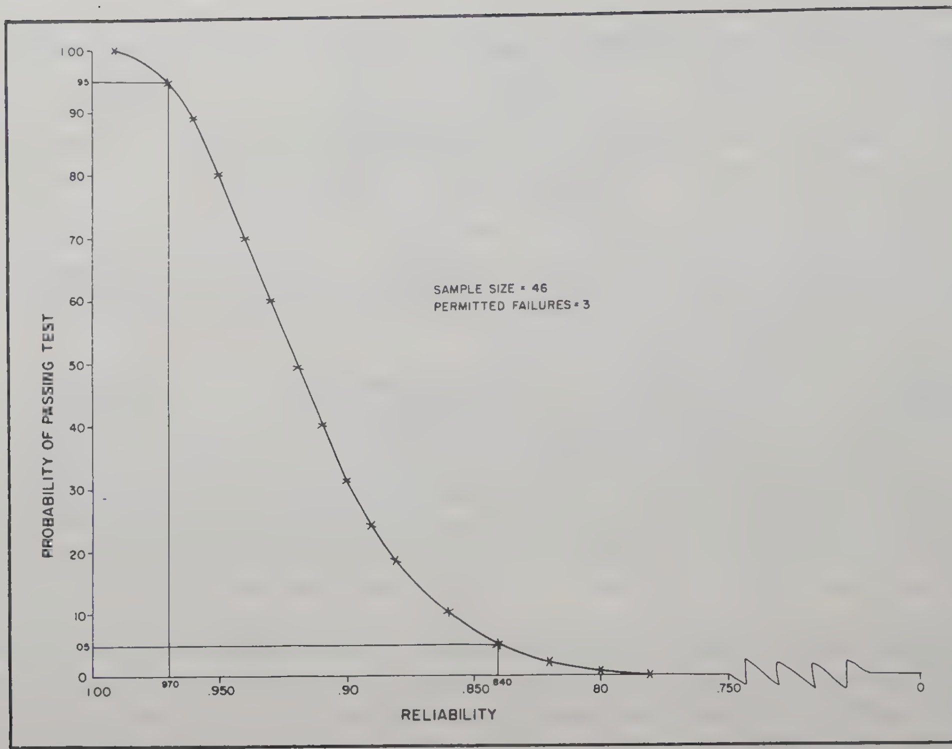

Fig. 1 - Operating characteristic curve.

the treatment it will receive in the hands of the customer. Therefore it is desirable to prepare an environmental chamber that will be capable of exactly matching the levels and variations of temperature, humidity, vibration, shock, etc., that the equipment will experience in actual flight. We shall place each of the units in this chamber and run it through a simulated flight. Then we shall take it out and perform the proper maintenance procedures on the unit, put it back in the chamber and run it again. This procedure will be repeated until we have reached the normal life expectancy of the equipment (i.e., the point at which it would be replaced in the aircraft). If a unit operates properly during each of the simulated missions in the chamber, we consider that it has passed the test; if it fails during any one of these missions, we consider that it has failed the test.

Stringent though the requirements of this test may appear, there is no alternative if we are going to qualify the equipment in a statistically satisfactory manner. Unfortunately, the test requirements which have been described are more than difficult to meet. They are, in most cases, completely impossible. A sample size as large as the one indicated may be very hard to obtain, particularly if we are producing small quantities of expensive and complex equipment. But this is the least of our troubles.

How do we know that the 46 units about to be tested are a homogeneous representation of a homogeneous population? In other words, how do we know that our sample is representative of what we will produce six months or six years from now? Even granted the proper number of units for the qualification, and that these units represent future production, we almost never know in advance the complete details of the equipment's operating environment. Even this knowledge, if

we had it, would not be a complete answer because we do not have test equipment which will accurately simulate anything but the most simple environmental conditions. And, then, to make matters worse, there is the question of time. The qualification procedures we have described would take months or perhaps years to complete, depending on the quantity of test equipment and manpower available. This means a time gap between pilot-line production and full-scale production that is simply not economically tolerable.

We are faced with a serious dilemma. We have, on the one hand, the very real necessity for the reliabiltty qualification of our equipment. And, on the other, we find that a statistically valid demonstration is beyond our means in time, knowledge, money and equipment. Obviously, unless we are willing to give up, some other way of tackling the problem must be found. I believe there is another way. But to find it, we must be willing to make some radical changes in our thinking. First, we must change some of the assumptions which are nearly always implicit in our thinking concerning the reliability of complex equipment. Second, we must combine engineering intuition and statistical facts into a shrewd maximization of whatever test results we have the means and the knowledge to obtain.

Some of the assumptions usually made are: that the statistically valid test discussed earlier makes sense only under one set of conditions; that our equipment is only barely up to the required reliability level; and that in order to check the fact that it has just squeaked by we must use a very delicate and exact reliability test. This is nonsense. No one knows how to build electronic equipment to a precise reliability specification. Even if they knew how, since when does good engineering practice consist of just barely squeaking by a specification? In our example, we were told that the reliability goal was 0.97. The only sensible way to interpret this is that every unit of equipment must operate when and where desired. We were also told that a reliability of below 0.84 is not acceptable. This simply means that we are still permitted to be human.

With this point of view, we tackle the job of designing equipment that will not fail under conditions that are significantly worse than anything the customer can serve up. When we are through, instead of assuming that our equipment is only just reliable enough and trying to devise the very difficult series of tests that will prove or disprove the point, we can assume that the equipment has a very comfortable reliability margin built in and then test to see if this reliability margin really exists.

Let us look at one or two of the advantages of this method of reliability qualification and then follow through a simplified example of its use. We have pointed out the difficulty of obtaining advance information of the environmental stresses our equipment will experience during its operation. A full statistical evaluation of the equipment's reliability requires considerable knowledge not only of the levels of these stresses but also of the expected variation of these stresses. If, on the other hand, we are qualifying by the reliability safety margin method, it is only necessary to obtain, or perhaps intelligently guess, the maximum levels of these stresses. Obviously, the latter information is more likely to be available than the former.

We can also show an improvement in the number of units necessary for reliability qualification. Naturally the more units available the better, but if we

are restricted to a fairly small sample size, the qualification can still be adequate. Because the basic design specifications have included reliability margins, we need not expect many units to fail under test. This will be particularly true if component part and subsystem qualification tests have been part of the reliability program. Also, except for life testing, there is no reason why the same units cannot be used for the different parts of the qualification test. For these reasons, the necessary sample size will rarely be as large as that necessary for the previously described statistical reliability evaluation.

For the sake of illustration, suppose we expect a maximum operating temperature of 150°F, and have set a reliability margin of 20°F for our stabilization system black box. The equipment, consequently, has been designed to operate at 170°F. Also suppose that 20 units have been made available for reliability qualification. We might take five of these units and test them at temperatures between 130°F and 170°F, the 170°F test having a time length equivalent to the expected mission. If all five units operate adequately under these temperature stresses, we need have little hesitation in qualifying the equipment for operation at 150°F. If one or more of the units fail, we have a problem. Our procedure now should be to determine statistically the probability that the unit will fail at 150°F. If this probability is judged too high, necessary design changes must be made. To obtain a better sample for the statistical check, it would be wise to run another five units through the temperature test. Figure 2

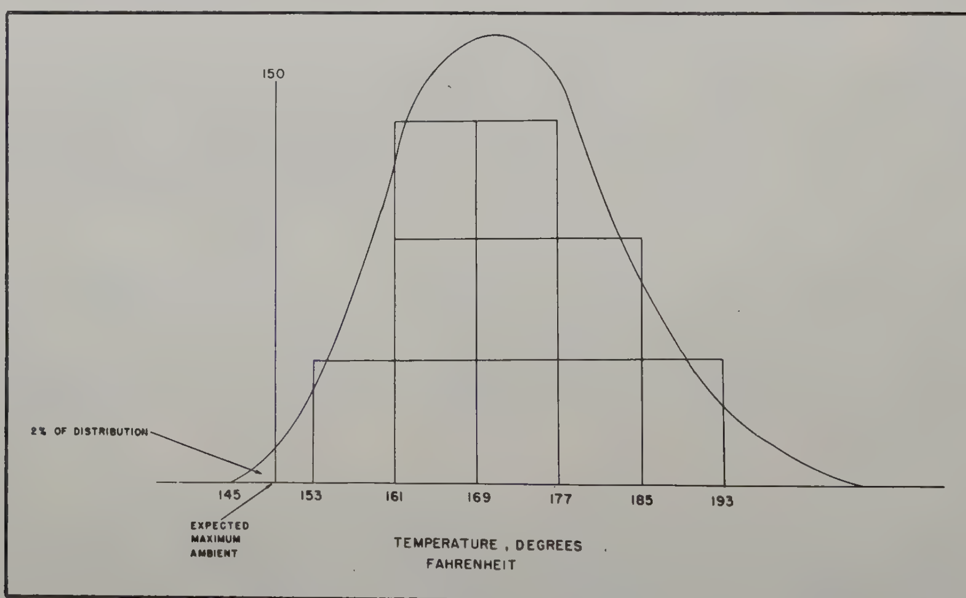

Fig. 2 - Fit of distribution to observed temperature failures.

illustrates the possible results of this investigation. Each of the boxes represents a unit which has failed in the indicated temperature range. A distribution has been fitted to this sample and it will be noted that the tail of this distribution overlaps the 150°F mark. It is estimated that in this case there is a 2 per cent probability that the equipment in general will fail at 150°F. Of course, this example has been simplified for the purposes of illustration. The situations which arise in actual practise tend to require considerable statistical sophistication for their analysis.

Thus, we have substituted a combination of engineering and statistics for a purely statistical reliability qualification of equipment. At the same time, we have concentrated our qualification procedures on proving the existence of comfortable reliability margins, rather than on proving the existence of specific reliability. In so doing, the whole process of reliability qualification has been brought down to the level of reasonably easy accomplishment.

Although it might appear on the surface that we have given ourselves a particularly difficult job by using a reliability safety margin philosophy of design and that we have been forced to overdesign and deliberately exceed our reliability requirements, any experienced reliability engineer knows that this is the only way very high reliability is ever achieved. It surely makes as little sense for us to avoid safety margins as it does for the builder of bridges to avoid them. And, it should be added, the wide-scale use of this method will do much to build up a badly needed series of electronic components specifically designed to operate with reliability at extreme environmental levels.

I have asked for a new look at the problem of the reliability qualification of electronic equipment. The task is not easy, even under the best circumstances. But, with the proper point of view, the job is at least possible. After these methods have been used for a while and the results correlated with the performance of the equipment in the hands of the customer, the procedures can be modified and refined. Eventually, as our knowledge and experience grows, we can substitute certainty for doubt, and send out our equipment with the knowledge that it has been designed for reliability, qualified for reliability, and ready to perform with reliability.

PASSIVE COMPONENTS FOR SUBMARINE TELEPHONE CABLE REPEATERS

M. C. Wooley
Bell Telephone Laboratories
Murray Hill, New Jersey

The most difficult problem in the design and production of passive components having an extreme degree of reliability is, in most cases, that of knowing when or to what degree the reliability goal is achieved. The second greatest problem is that of foreseeing or predicting in what respects each class or type of component is most likely to fail so that appropriate counter measures can be taken. This paper will discuss each of these problems as it applies to capacitors, inductors, resistors and transformers as designed and produced for use in submarine cable amplifiers.

It is only fair to say that we do not have the answer to the first problem, in so far as the transatlantic submarine cable system components are concerned. There are, in the deep sea portion of the transatlantic telephone cable system, approximately 6,000 passive components for which the goal is no failure in 20 years. If we are to be 90 per cent certain of achieving this degree of reliability, the average annual failure rate must be of the order of 1 in 1,000,000 or less. Obviously no practical program of testing can hope to detect this small failure rate. In fact, it is estimated that we would have to run tests on 6,000 components for more than 400 years in order to obtain sufficient data to permit an estimate of such a low failure rate.

Lacking the ability to determine from tests what the failure rate may be, let us examine what can be done to insure the lowest failure rate practicable. With the exception of the high-voltage paper capacitors in the submarine cable repeaters, the passive components do not "wear out." True, they may age or drift in value, but, aside from drift or aging, the most likely failures are the catastrophic ones. Catastrophic failures may be due to poor electrical connections, broken conductors, corrosion which may result in open or short circuits, chemically unstable materials which may give off corrosive products, mechanically unstable materials which may cold flow or break and, finally, foreign material which may cause short or open circuits. Since all varieties of a given type of component are not subject to the same defects nor to the same degree, the first step was to select those varieties of components which would do the job at hand and at the same time be least subject to defects.

Because reliability in any complex system such as the telephone plant is always important, statistics on the performance of many varieties and types of components are available to us. From this information those types which had the best record of trouble-free service were selected as candidates for submarine cable repeater use. This approach is restrictive in that it rules out a number of promising types simply because they are new. It does, however, provide a firm background on which, with extra care in design and manufacture, an extremely reliable series of components can be based. This procedure led to the use of wire-wound resistors, impregnated paper and silvered mica capacitors, molybdenum permalloy powder cores for those inductors requiring magnetic cores and molybdenum permalloy tape cores in transformers. The use of ferrites for magnetic cores,

carbon films in resistors and plastic films in capacitors were all ruled out on the grounds of lack of proven long-time stability or reliability.

It is appropriate at this point to discuss briefly the importance of stability in a system involving many amplifiers in tandem and their being inaccessible for adjustment after installation. In such a system the margins are such that a change of a few db in gain for the over-all system can seriously degrade performance in so far as noise and overload are concerned. With 51 repeaters or amplifiers in tandem this means that any systematic change or deviation from the correct value for a given component must be so small that it will not produce more than a few hundredths of a db change in the gain characteristic of each repeater.

Components in some circuit positions are more critical than in others so that initial tolerances range from ± 3/4% to ± 7% for paper capacitors, from ± 0.25% to ± 1.0% for mica capacitors, ± 0.1% to ± 1.0% for resistors and ± 0.25% to ± 1.0% for inductors. Requirements on stability, as measured by the change taking place during temperature cycling, were generally of the order of 1/5 to 1/10 of the initial tolerances. Consequently, when we are considering reliability, stability as well as freedom from complete failure must be considered. Having thus established the types and some general requirements for the components, let us examine them in more detail.

DESCRIPTION OF REPEATER

The repeater itself is made up of a total of 17 methyl methacrylate sections each approximately 5 inches long and 1-3/4 inches in diameter. These are mechanically coupled end to end by springs so that the over-all length of the repeater is approximately 8 feet. Each section consists of a double-walled cylinder which contains a plastic core provided with recesses to receive the components. In the case of large capacitors and the electron tubes, a single component is contained in a section. Flat bus tapes which provide the electrical connections between sections are laid in grooves between the two cylinders. The complete repeater is housed in a double layer of close fitting steel rings, and this in turn in a copper tube with a 1/32 inch wall. With this arrangement the components of the repeater are limited to a maximum of 1-3/16 inches in diameter by 4-1/16 inches long.

Electrically, the repeaters in each cable are all connected in series; i.e., the heaters of the 3 tubes in each repeater are in series with each other and with all other repeaters. Plate voltage is supplied by the voltage drop across the 3 heaters in each repeater. Power, at constant current, is fed to each end of the cable at approximately 2,000 volts dc. Consequently the voltage to ground drops gradually to zero as we progress along the cable to the center of the span. Two cables are used, one for transmission in each direction.

INDUCTORS

Because of the proximity of the steel rings which protect the amplifier from sea bottom pressure, it was necessary to use closed cores for all inductors. Consequently they were all wound in toroidal form using either nonmagnetic or molybdenum permalloy dust cores. Some were wound with resistance wire to save the space which a separate resistor would require and also to eliminate a joint.

Both inductance and resistance of such coils were controlled to close limits by providing separate adjustments. Inductance was adjusted by removing turns, and resistance by removing wire from a "noninductive" winding. In one case a magnetic core of larger cross-section than could be fitted into the container with a conventional toroid was required. In this, two cores arranged in a figure-eight formation were threaded by the same winding. Typical examples of the inductors are shown in Fig. 1.

The hazards associated with inductors are: (1) broken wire or terminal leads resulting from abnormal flexing; (2) shorted turns; and (3) mechanical instability. Since every soldered joint is considered a potential source of trouble, the repeaters and components were designed to minimize the number of such connections. This meant that the inductors had to be made without joints within the winding, the wire of the winding being used to connect the coil into the circuit. With this arrangement the handling associated with adjusting and testing could result in the leads being flexed nearly to the point of breakage. Two procedures were used to prevent this. Where possible, the windings were arranged so that both ends were on the outside; the initial adjustment was made so that as a last operation an additional turn or turns could be removed to provide a lead of relatively unflexed wire. When this was not feasible, special handling fixtures were used to hold the inductor or transformer and its leads in a fixed relation to each other until it was ready for use. The greatest hazard for shorted turns occurs in multilayer windings or from crossovers in single layer winding, since these result in high pressure between turns. These hazards were minimized by carefully inspecting each layer of a winding for crossovers and by providing additional insulation between layers. Although Formvar enameled wire was used in most inductors, textile insulation as well as enamel was used in a few others.

Mechanical instability, either in the form of cracking or flow of the core, results in unstable inductance. To guard against this, nonmagnetic cores were

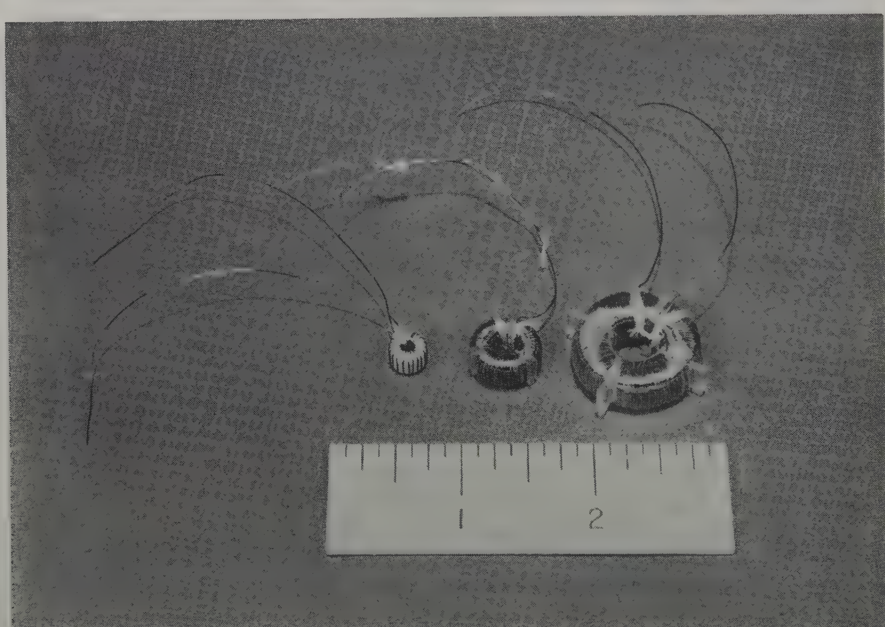

Fig. 1 - Typical toroidal inductors.

properly annealed and all cores were carefully inspected visually for cracks. The completed coils were also subjected to repeated temperature cycles and observed for stability. This was found to be a sensitive and effective control for such defects.

TRANSFORMERS

Transformers were used to couple the repeater to the cable. With the exception of their physical construction, they were of conventional design, using a wound molybdenum permalloy tape core and a spool-supported coil as shown in Fig. 2. Since a considerable part of the required gain-frequency characteristic of the repeater is obtained in the input and output networks, close control of

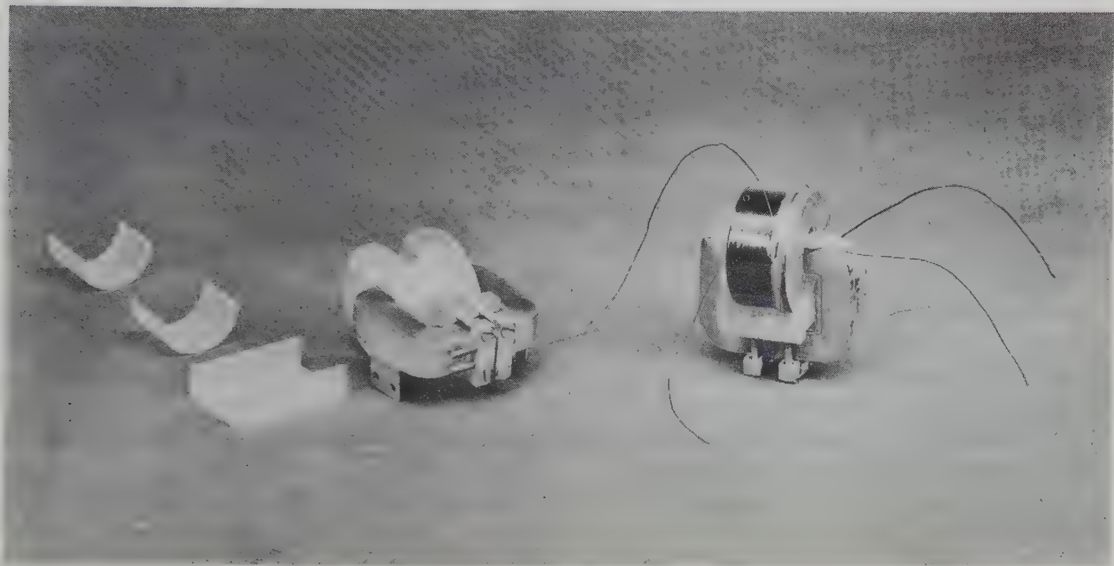

Fig. 2 - Transformer and its parts.

transformer parasitics, such as leakage and capacitance, was necessary. This was accomplished by a coil design which placed the windings in a fixed relation to each other and provided a high order of uniformity throughout the product. A precisely adjusted and stable air gap was also required. This was assured by applying a test winding to the core before the regular winding and observing the stability during mechanical stressing and temperature cycling. In general the same factors apply to the reliability of transformers as to inductors and likewise the same precautions must be taken to insure reliability.

Although all of the resistors for the repeater were wire-wound types, they were of many sizes and shapes to meet the physical and electrical requirements of the repeater. Some were simple inductive single or multiple layer windings on appropriate forms, while others were windings of mandrelated wire; i.e., wire wound on a silk core and protected by a textile serving. The most critical part of their construction was the terminal connection. The wire size was limited to #46 gauge and larger in the interest of reliability. However, since #46 is only 1-1/2 mils in diameter, it is very susceptible to breakage. Stranded lead wires were attached to the fine resistor wires by brazing. The processing of such splices was very carefully controlled so as to insure a good electrical connection, to avoid flexing or overheating the resistance wire adjacent to the splice.

The production of satisfactory splices was one of the most critical of all component manufacturing operations and normally required several weeks of operator training before acceptable splices could be made. Figure 3 illustrates the steps in making and protecting such a splice.

In addition to combined inductors and resistors, resistors were also included in the same container with some of the paper capacitors. This was done only for space reasons as they have nothing in common except circuit positions. In this case the resistors were constructed of materials which would withstand capacitor drying and impregnation processes and at the same time not contaminate the capacitor. Consequently the resistors used ceramic spools, enameled wire and capacitor paper as interlayer insulation. These and other examples of the resistors are shown in Fig. 4.

In spite of the fact that in all components extreme care was taken to use materials which were compatible, in two instances, both involving resistors, the hazards of bringing together two new materials were encountered. Most of the metal parts of components were gold-plated to improve solderability and to avoid the growth of metal whiskers. This included the lead wire used on some components. In those designs in which the resistor was included in the same container with a capacitor, these lead wires were threaded through small holes in the resistor spool. In early models phenol fiber spools were used, but for actual use in the repeaters a less chemically active material was desired for use inside the capacitors. When a ceramic was substituted for the fiber, the sharp edges around the holes acted as knives and scraped long fine slivers of gold from the resistor lead wires. Such slivers were certainly undesirable additions to a capacitor in which clearances between uninsulated parts were of the order of 3/32 inch. Only by careful inspection were the slivers detected originally, and they were eliminated only by rounding the edges of the holes and careful assembly of the parts.

The second instance was perhaps less hazardous but equally unexpected. To prevent contamination it was specified that only new or carefully cleaned supply

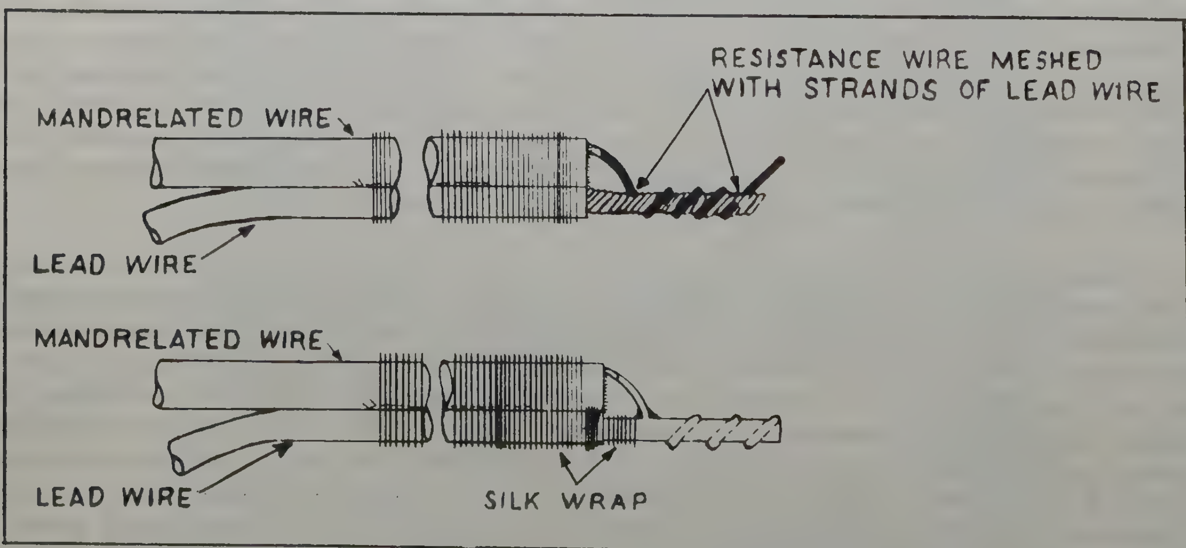

Fig. 3 - Formation of fine resistor wire - lead wire splices.

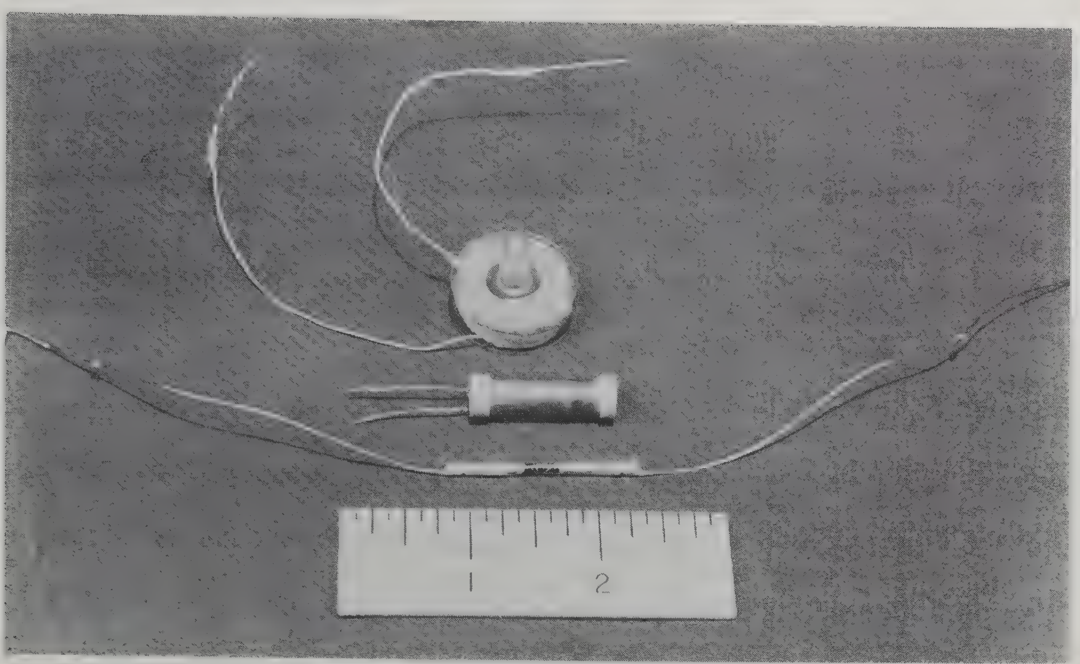

Fig. 4 - Wire wound resistors.

spools should be used for wire. For the mandrelated wire the spools were cleaned initially with steel wool and, as it turned out, bits of this remained throughout subsequent inspections. A sharp-eyed inspector found a bit of steel imbedded in the nylon serving of the mandrelated wire. As a result, all of the wire used was passed through a sensitive, magnetic particle detector. A few other particles were found, but only one of these in a critical resistor could have resulted in a short or open and failure of the system.

CAPACITORS

About fifteen years before the laying of the transatlantic cable, work was started on the development of the high-voltage capacitors for such a system. The space and electrical requirements for these capacitors required that they contain more capacitance per unit volume than most commercial designs. Therefore, a fairly extensive program of life testing of various combinations of papers, foils and impregnants was conducted at both sea-bottom and room temperatures. In a fairly short time this program showed the superiority of liquid impregnants at low temperatures and ultimately led to the use of a castor-oil-impregnated kraft paper-aluminum foil design.

Since this is the one passive component in the repeaters which does "wear out," a considerable amount of testing has been done to provide information from which a statistical estimate of the minimum life under service conditions could be made. The details of making this estimate have been published elsewhere,* so it will suffice here to outline it only briefly. From the total exposure of the life-test samples, in terms of capacitor-years at the maximum service voltage, during which only one sample has failed, we estimate by probability equations the range in time within which the first failure will occur in the system. This is

*"Flexible Repeater Design." Bell System Technical Journal, January, 1957.

analogous to estimating the per cent defective in a given sample when it is known that in another sample from the same universe a certain per cent is defective.

In our case the two samples did not, of course, come from the same universe. However, general experience as well as highly accelerated tests show that the life of the present universe is longer than that which the original samples represented. This is due mainly to improvements in capacitor paper during the past fifteen years as well as to improved control of processes and materials. Consequently, estimates of the life of the capacitors in the cable tend to be conservative. This estimate is dependent upon the number of capacitors in service and their service voltage. Since the voltage varies from repeater to repeater, it is necessary to translate the total exposure of the capacitors in the cable with their respective service voltages into an equivalent exposure of a smaller number of capacitors at the maximum service voltage. This is done with the so-called fifth-power rule, which states that the life of a paper capacitor is approximately proportional to the inverse of the fifth power of the applied voltage; i.e.,

$$\frac{L_1}{L_2} = \left(\frac{V_2}{V_1}\right)^n$$

where n ranges from 4 to 6.

From this we calculate that in one year the 306 high-voltage capacitors in the deep-sea portion of the transatlantic cable accumulate an exposure which is equivalent to that of 62 capacitors at the maximum voltage for the same length of time. Using this and the procedure outlined above, we have estimated with a probability of being correct 9 times in 10, that the first failure in the transatlantic system will occur in not less than 16 years nor more than 600 years.

There is the possibility of a catastrophic failure, perhaps the greatest hazard even in capacitors. Our best protection against failures of this nature is careful, unhurried construction by well-trained operators, supplemented by thorough inspection. For example, high-voltage capacitor units were wound at the rate of approximately 10 units per operator per day. This allowed time for the operator to observe irregularities in the materials or process. Many were found, including the remains of insects calendered into the capacitor paper.

The construction of both paper and mica capacitors followed conventional lines with some deviations in the interest of improved reliability. For example, the tension on the paper during winding was held within fairly close limits. Furthermore, its moisture content at the time of winding was controlled to facilitate meeting close capacitance limits and to minimize the spread in the capacitance aging from capacitor to capacitor. Also when unusually close capacitance limits had to be met, one electrode was made both narrower and shorter than the other, so that capacitance variations due to electrode misalignment were avoided. The paper capacitors all used "laid-in" or tab terminal construction, and in the high-voltage type these terminals extended through the terminal plate and served as the external terminals.

Two features of the silvered mica capacitors were unusual. First, contact to the silvered surfaces was made by interleaving fine silver foil between the laminations and clamping the ends of the laminations and the foil together by

the terminals. As a further precaution, the foil is soldered to the terminals so that the only dry contact is from silver to silver, and this covers a relatively large area. Second, the mica capacitors were simply mounted on a small slab of methacrylate by means of their terminals. This open type construction was used for all components except oil-impregnated capacitors because it avoids possible damage due to operations of housing or mechanical strains resulting from the housing. However, it does make more difficult the handling of components before and during their assembly into networks. Such construction is possible

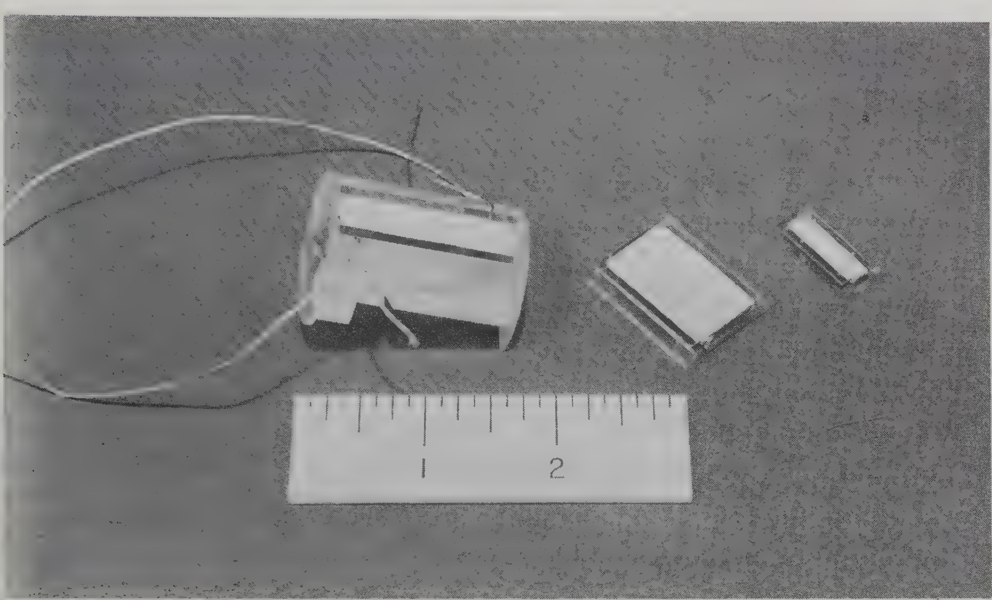

Fig. 5 - Silvered mica capacitors. Left - high surge voltage design.

because the repeater is thoroughly dried and filled with dry nitrogen before it is sealed. The open structure of the mica capacitors is illustrated in Fig. 5.

MANUFACTURING FACILITIES

A detailed description of the manufacturing facilities is beyond the scope of this paper, but it should be pointed out that the watchword was cleanliness. The whole manufacturing area was air-conditioned, and the air was cleaned with both mechanical and electrostatic filters. Operators wore special clothing to minimize lint, and the floors and work benches were regularly damp-cleaned. Special care was taken to avoid the accumulation of scrap materials such as small bits of fine wire, foil, filings, and the like, which could stick to hands or clothing and turn up in the wrong place. The paper-capacitor winding room was particularly restricted. Only winding and the assembly operations prior to impregnation were permitted in this area. The capacitor-impregnation area was also separated from other areas and was maintained at a slightly negative atmospheric pressure with respect to the other areas when oils or solvents were in use.

All machining of metal or plastic parts was done in areas isolated from those of fabrication and assembly, and all materials were inspected and cleaned before being brought into the assembly areas. Gloves, tweezers and vacuum pickup

tools were used extensively for handling parts and components, although the temperature and relative humidity of the manufacturing area was controlled to minimize perspiration. An exception was that paper-capacitor and resistor-winding operators were allowed to use their bare fingers for greater dexterity. Washing of hands was mandatory, whenever an operator returned from outside the working area, and in critical areas workers were encouraged to wash more frequently.

INSPECTION

Reference has been made several times to the critical inspection procedures used in the production of components, but it is difficult to convey in a few words an idea of the extent of this inspection. One way is to state that there was an inspector for every two production workers, and, in addition, that production people were trained to inspect their own product. Consequently, they turned over to inspection only those products which they felt would pass inspection. Another and perhaps better mode of illustration would be to take a typical component and list the steps in its production and inspection. For this purpose, let us take a silvered mica capacitor which is neither the simplest nor most complicated from an inspection standpoint.

However, it should be pointed out that the inspection started with a thorough examination of the raw materials. The specifications on these ranged from standard ASTM designations to specialized and elaborate requirements intended only for submarine cable use. In all cases, however, the sampling rate was much higher than usual. For example, capacitor paper was sampled at the rate of 1 test sample for each 3 pounds of paper, and each sheet of methyl methacrylate was sampled. Parts for components were 100 per cent inspected, and wherever it was applicable (chiefly with plated metal and ceramic parts) a water extract conductivity test was used to insure that they were free from contaminants such as plating salts or perspiration.

Returning to the example of component inspection, the major steps in the production and inspection of a silvered mica capacitor are shown in Table I. The inspectors were trained to be on the lookout at all times for all types of defects, including those covered by previous inspections. In addition, much of the visual inspection was done with the aid of low-power binocular microscopes.

All data taken were recorded and initialed by the inspector, and the results of each inspection were likewise recorded and initialed for each individual component. For this reason, as well as to make it possible to trace the history of each component, they were assigned individual serial numbers. This system was extended into the inspection of raw materials which were identified by an appropriate numbering system. Consequently, as part of the final approval for the use of a component, its history was traced to insure that all the raw materials used in it were inspected and had met their requirements and that all the specified operations and inspections on the component itself had been carried out. The lack of any part of this data caused the component to be rejected.

One might reasonably ask how much such detailed inspection contributes to improved reliability. Certainly no numerical value can be attached to it. However, a list of some of the things uncovered, which would probably not have been found in normal manufacturing procedures, is some indication of its value. I

TABLE I

Manufacturing and Inspection Operations
for Silvered Mica Capacitors

After cleaning mica laminations inspect for:

Foreign matter
Spots and stains
Air inclusions
Scratches, cracks, holes and delamination

After silvering inspect for:

Silver thickness
Uniformity
Stains and foreign inclusions
Mechanical damage
Dimension of silvered areas
Capacitance

After stacking inspect for:

Alignment
Proper soldering
Mechanical damage
Capacitance and conductance

After mounting inspect for:

Proper cementing
Freedom from cement on or between laminations
Mechanical damage
Capacitance, conductance and dielectric strength

After life test for 4 months at 400 v inspect for:

Capacitance, conductance and dielectric strength
Change in capacitance during life test
Mechanical damage
Dirt or contamination

---

have already mentioned insect remains in capacitor paper and steel-wool splinters in mandrelated wire. Other undesirable factors are:

(1) Pieces of nickel plating which had flaked off tweezers used in assembly and forming operations.

(2) Floor sweepings in a roll of capacitor foil.

(3) Damaged wire.

(4) Mechanical damage from misaligned winding machines.

(5) Damaged splices in resistors.

(6) Winding errors found by measurement of the parasitics of each resistor.

(7) Oil leaks in capacitors not disclosed by normal leak tests but revealed by extensive temperature cycling.

(8) Loose solder in capacitors found by X-ray examination.

(9) Inadequate impregnation of a lot of capacitors disclosed by a destructive sampling life test on each impregnation lot.

(10) Cracked cores in inductors found by temperature cycling.

While some of these would not necessarily be disastrous, many of them would be capable of causing a failure of the system. If only one failure is prevented by the elaborate inspection, the cost of the extra care has been more than justified.

STATISTICAL ASPECTS OF RELIABILITY IN SYSTEMS DEVELOPMENT*

John S. Youtcheff
Missile and Ordnance Systems Department
General Electric Company
Philadelphia, Pennsylvania

Summary -- A system reliability program is an operational procedure for obtaining the over-all reliability objectives in the development of a system. The reliability program must be closely integrated with the system developmental activities to assure that the over-all program objectives are fully obtained within the required time scale. The analytical objectives of the system reliability program are twofold. First, the specified system reliability requirements must be apportioned to the subsystems and components to assure adequate equipment design. Second, as component and subsystem design data are made available throughout the development program, this information must be utilized in predicting the system reliability. In predicting system reliability, it is necessary to determine both the component reliability relationships and the individual component failure probabilities. Several statistical methods for determining component reliability are presented; however, the exact methodology must be tailored to the specific system and development program. System reliability can be predicted by utilizing component reliability data together with an adequate analysis of component and subsystem reliability relationships.

INTRODUCTION

The problem of obtaining system reliability in a development program is extremely complex. There is no single reliability program or methodology that is applicable to all systems. As a result, it becomes necessary to review the possible analytical methods available in the field of reliability and to suggest their use in obtaining the objectives of a particular system development program. It is the intent of this paper to discuss the reliability problem, system program considerations and possible statistical methods of attack. This approach can cover only a limited aspect of the over-all problem presented here.

THE RELIABILITY PROBLEM

Reliability is the probability that a device will perform its required function under given environmental conditions for a specified operating time within the prescribed limits of precision and accuracy. A device is any mechanism from a single part or component in a defined operational system to the system itself. A failure is defined as performance of a device outside specified operating limits. Under this definition, a failure may be the result of performance degradation or complete inoperation. System unreliability may result from either component failures or from two or more components deviating from their normal mode of operation to cause system failure by their combined effect.

The three major types of failures (Fig. 1) are initial failures, chance failures and wear-out failures. If a device fails at the beginning of use, this

*This paper was presented at the Sixth Annual Symposium on Statistical Methods in New York, New York, November 29, 1956.

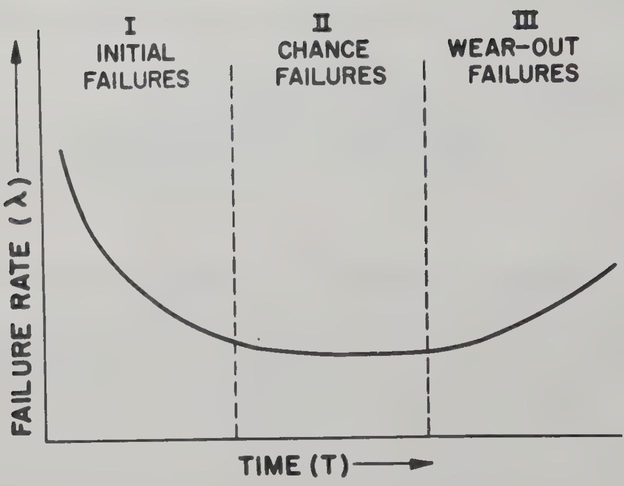

Fig. 1 - Modes of failure.

$$P_n = \frac{(t/\bar{T})^n \, e^{-t/\bar{T}}}{n\,!}$$

WHERE P_n IS THE PROBABILITY OF HAVING "n" FAILURES IN TIME "t", AND $\bar{T}$ IS THE MEAN TIME BETWEEN FAILURES.

Fig. 2 - Chance-failure equation.

failure is termed an initial failure. Initial failure is caused by either poor design or an adverse environmental effect prior to use. It should be noted that poor design may result from either unsatisfactory engineering or a defective unit in a well engineered lot.

If a device fails under an unexpected environmental condition, which is too severe or not anticipated, this failure is termed chance failure. This results from the operational load exceeding the initial strength of the component.

If a device fails as a result of changes in a significant characteristic throughout the life of the device, this failure is termed wear-out failure. This change can be expressed in terms of a deterioration rate acting on the initial strength of the component. Wear-out failure occurs when the strength of the device, as a result of extended deterioration, falls below the operational load.

This classification of failures and this pattern with respect to operating time provides a means for the quantitative prediction of equipment failures. The probability of having a given number of failures (zero, one or more) in a specified time interval is equated in Fig. 2. This equation provides a basis for a time-to-failure analysis.

The choice of methods and techniques in reliability analysis, however, depends upon the reliability problem or, more broadly, the reliability objective or specification requirements. Specification of reliability requirements is a

problem in system procurement. The type of system to be procured determines the specified reliability requirement, and the variance in systems is matched by the variance in reliability specifications.

A common form of reliability specification is that in which a set of reliability and corresponding confidence values P_1, P_2, α and β are given, where these specified values determine the statistical details of a design acceptance test. The design acceptance test is concerned with whether or not the reliability requirements have been met. An example of such a reliability specification is shown in Fig. 3 and further illustrated by the operating characteristic curve in Fig. 4 which relates system reliability to probability of acceptance. In Fig. 4 the specified values P_1, P_2, α and β determine the shape of the curve, the sample size to be tested N and the acceptance number C. The requirements illustrated in

P_1	(MINIMUM ACCEPTABLE RELIABILITY) (FOR TIME PERIOD TO)	= .70
P_2	(RELIABILITY DESIGN OBJECTIVE)	= .90
α	(PROBABILITY OF ACCEPTING P_1 OR LOWER)	= .10
$1-\beta$	(PROBABILITY OF REJECTING P_2 OR HIGHER)	= .10

Fig. 3 - Reliability requirements.

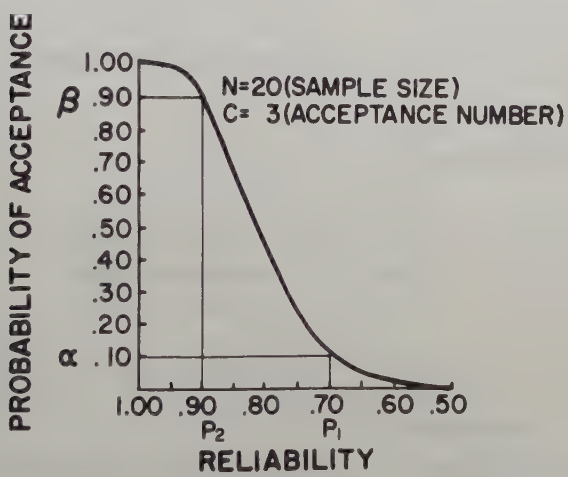

Fig. 4 - Operating-characteristic curve.

Fig. 4 afford 90 per cent protection or confidence against accepting a design as poor as 0.70 and afford similar protection against rejecting a design as good as 0.90. Again, in the example of Fig. 4, twenty complete systems tests should be conducted and the design accepted if three or fewer failures occur. It is thus seen that the design acceptance test based on the numerical values established in the specification requirements affords a method for procuring a system of the desired reliability.

The reliability problem in systems development is concerned with attaining the reliability requirements specified for the system. The program and methodology for attacking the reliability problem are discussed in the following sections.

THE SYSTEM RELIABILITY PROGRAM

It is the purpose of the system reliability program to establish an operational procedure for arriving at an optimum solution to the reliability problem. It should be derived by analyzing the system developmental activities and integrating the various reliability functions and considerations necessary to obtain the over-all reliability requirements. An outline of the developmental phases in a systems development program is shown in Fig. 5. It should be emphasized that the reliability program is not directly concerned with all of the steps outlined in Fig. 5; however, the results of these steps are necessary for an adequate

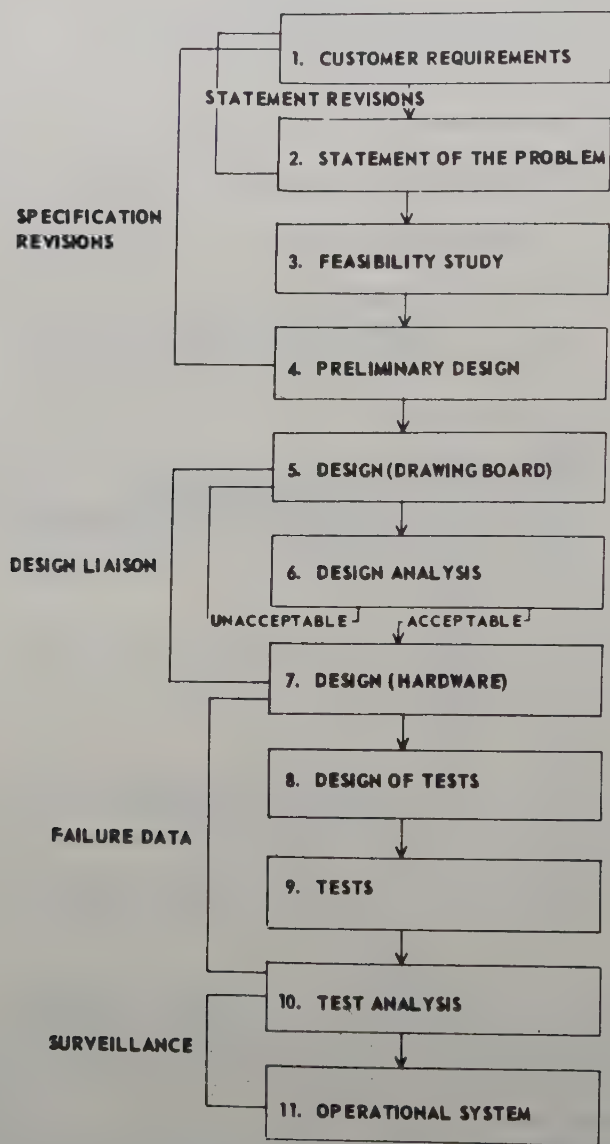

Fig. 5 - Phases of a system development program.

analytical approach to the reliability problem. This procedure provides a flexible mechanism so that all required data will flow in a logical sequence allowing for all necessary feedback.

The reliability program itself consists of a series of phases linked to salient system developmental steps. This linkage provides a means for obtaining the necessary data inputs required for the reliability analysis, or model, as well as providing for the efficient handling of the required data outputs (Fig. 6). Although the reliability program is essentially concerned with the necessary reliability analysis, it should fully provide for the required liaison, data collection and evaluation activities. The analytical objectives of the system reliability program are twofold. First, the specified system reliability requirements must be apportioned to the subsystems and components to assure adequate equipment design. Second, as component and subsystem design data are made available throughout the development program, this information must be utilized in predicting the system reliability (Fig. 7). Quantitative reliability data for

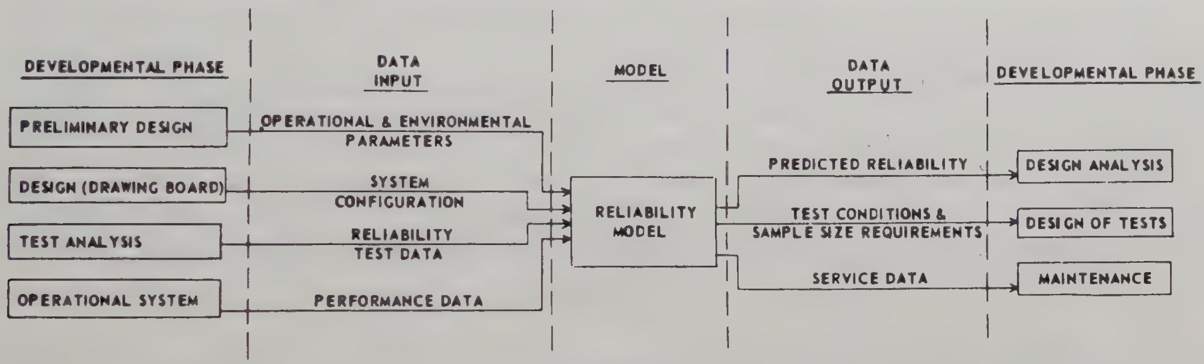

Fig. 6 - System reliability model.

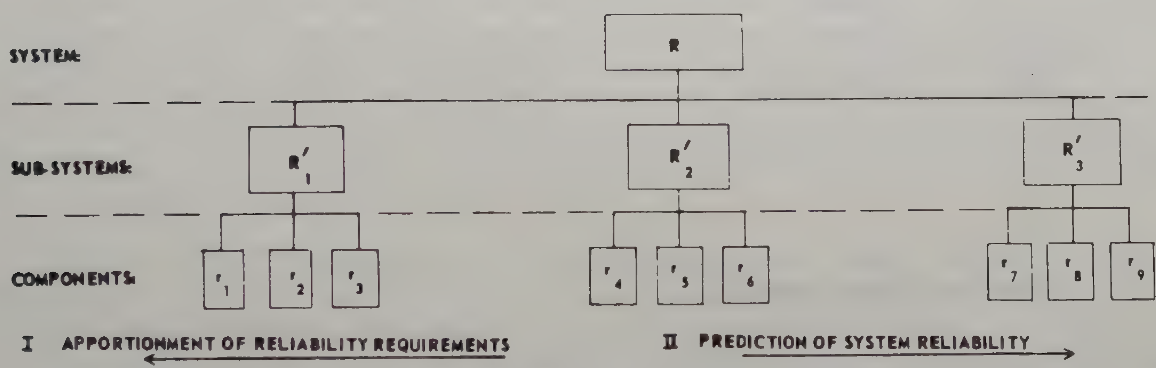

Fig. 7 - The system reliability problem.

the system, subsystems and components should be continuously supplied through an iterative process, extending throughout the program from initial reliability apportionment to final system reliability prediction.

RELIABILITY METHODOLOGY

In apportioning the system reliability requirements to the components of the system, and in predicting the system reliability from component reliability data,

both functional and operational factors should be considered. The functional dependency between the components is here considered in the form of series and parallel type structural arrangements. A series type component is part of a chain of components in which the failure of one component results in the failure of the system. If it is assumed that these components in series are mutually independent, the reliability of the system is equal to the product of the individual component reliabilities. A parallel type component is part of a redundant network of components in which the failure of all like components is necessary before system failure occurs. If it is again assumed that these components in parallel are mutually independent, and if the failure of a component always produces a system failure, the reliability of the system is equal to one minus the product of the individual component failure probabilities. Considering only functional dependency, the general reliability equations are presented in Fig. 8. The effects of performance dependency can be obtained by proper performance analysis, and the effects of failure dependency can be obtained through adequate correlation of component test data or from an analysis of systems test results.

In predicting system reliability, it is necessary to determine both the component reliability relationships and the individual component failure probabilities. In a development program, many (if not the majority) of the components have little or no performance or reliability history. Although component reliability can be determined by adequate test data, careful consideration must be given to a realistic sampling plan. A realistic sampling plan is here defined as one that is optimally balanced for data output within cost and time limitations. In Fig. 9 two basic sampling methods are compared. Sampling by attributes

SERIES ARRANGEMENT:

$$R = \prod_{i}^{m} r_i = \prod_{i}^{m} (1 - q_i) \tag{1}$$

PARALLEL ARRANGEMENT:

$$R = 1 - \prod_{i}^{n} (1 - r_i) = 1 - \prod_{i}^{n} q_i \tag{2}$$

SERIES - PARALLEL ARRANGEMENT:

$$R = \left[\prod_{i}^{m} (1 - q_i) \right] \left[1 - \prod_{i}^{m} q_i \right] \tag{3}$$

WHERE R IS THE RELIABILITY OF THE SYSTEM,
q_i IS THE FAILURE PROBABILITY OF "m" SERIES COMPONENTS,
AND q_i IS THE FAILURE PROBABILITY OF "n" PARALLEL COMPONENTS

Fig. 8 - Reliability prediction equation.

I SAMPLING BY ATTRIBUTES:	
ADVANTAGE:	RIGOROUS SOLUTION.
ASSUMPTION:	RANDOMNESS OF SAMPLE.
INFORMATION REQUIREMENT:	FRACTION DEFECTIVE.
CALCULATE:	RIGOROUS PROBABLE LIMITS WITHIN WHICH THE OBSERVED SAMPLE FRACTION DEFECTIVE SHOULD LIE.
II SAMPLING BY VARIABLES:	
ADVANTAGE:	SMALLER SAMPLE SIZE REQUIREMENTS, IN GENERAL.
ASSUMPTION:	FUNCTIONAL FORM OF UNIVERSE.
INFORMATION REQUIREMENT:	MEAN POINT OF FAILURE AND STANDARD DEVIATION.
CALCULATE:	LIMITING PROBABILITIES THAT STATISTIC LIES WITHIN CERTAIN BOUNDS.

Fig. 9 - Establishing component reliability.

has the advantage of resulting in a rigorous solution but usually requires a very large sample size to verify a component's reliability which is close to unity. This large sample size requirement usually imposes an excessive (if not prohibitive) burden on the time and cost of the system development program. Sampling by variables has the advantage of generally requiring a smaller sample size, but the disadvantage of usually resulting in a less rigorous solution than the attribute sampling method. Sampling by variables, however, does have the advantages of a low cost compared with other methods, of being capable of accomplishment early in the program and of being able to uncover main design weaknesses while giving statistical results.

It should again be noted that the optimum sampling technique can only be selected after careful consideration is given to the particular component, system and program requirements. Average sampling varies with (1) risk (the lower the α and β values, the more testing is required); (2) required reliability (the smaller the P_2-P_1 value, the more testing is required); and (3) reliability range (the closer to unity, the more testing is required).

Sampling by Attributes

A system consisting of m components in a series arrangement will now be considered. The best estimate of component reliability can be expressed as the ratio of the number of successful outcomes to the total number of tests made on the i th component. The best unbiased estimate of the system reliability can be expressed as the product of the component reliabilities. The β per cent lower

bound, R_L (giving β per cent confidence that $R_L \leq R$) can also be determined. These relationships are equated in Fig. 10.

A system consisting of n components in a parallel arrangement is also considered. The best estimate of component reliability can again be expressed as the ratio of the number of successful outcomes to the total number of tests made on the j th component. The best unbiased estimate of the system reliability can be expressed as one minus the product of the component failure probabilities. The β per cent lower bound, R_L, is also determined for this case. These relationships are equated in Fig. 11. The value t_β is defined by the normal probability integral. Values of t_β corresponding to selected values of β are given in the table in Fig. 12 along with the normal probability integral.

Because of time and cost considerations it is desirable to discontinue sampling as soon as the reliability of the equipment has been verified to the desired confidence. To do this, it is necessary to know when sufficient samples have been observed to prove or disprove the reliability hypothesis. A sequential sampling plan, utilizing the specified values of P_1, P_2, α and β as equated in Fig. 13, is most suitable for this purpose. A graph can be prepared with one axis representing the number of successful tests (gm) and the other axis representing the number of unsuccessful tests (dm). Bounds are imposed, within the graph, by Eqs. (1) and (2) of Fig. 13. These bounds establish action areas for accepting or rejecting the equipment. This graph is shown in Fig. 14. In the sequential sampling, the actual series of tests will be represented by a stepped line joining the series of points. Any given stage in the sampling process will

SERIES ARRANGEMENT:

$$\hat{r}_i = \frac{n}{N_i} \tag{1}$$

WHERE $\hat{r}_i$ IS THE BEST ESTIMATE OF COMPONENT RELIABILITY,
n_i IS THE NUMBER OF SUCCESSFUL TESTS,
AND N_i IS THE TOTAL NUMBER OF TESTS.

$$\hat{R} = \prod_i^m \hat{r}_i \tag{2}$$

WHERE $\hat{R}$ IS THE BEST ESTIMATE OF SYSTEM RELIABILITY.

$$R_L = \hat{R} - t_\beta \sqrt{\prod_i^m \left(\left[\hat{r}_i + (N_i - 1)\,\hat{r}_i^{\,2}\right]/N_i\right) - \hat{R}^2} \tag{3}$$

WHERE R_L IS THE β% LOWER BOUND, DETERMINED BY t_β

Fig. 10 - Determination of the β% lower bound on system reliability (series arrangement).

PARALLEL ARRANGEMENT:

$$\hat{r}_i = \frac{n_i}{N_i} \qquad \text{OR} \qquad \hat{q}_i = 1 - \frac{n_i}{N_i} \tag{1}$$

WHERE $\hat{r}_i$ IS THE BEST ESTIMATE OF COMPONENT RELIABILITY,
$\hat{q}_i$ IS THE BEST ESTIMATE OF COMPONENT PROBABILITY OF FAILURE,
n_i IS THE NUMBER OF SUCCESSFUL TESTS,
AND N_i IS THE TOTAL NUMBER OF TESTS

$$\hat{R} = 1 - \prod_i^n \hat{q}_i \tag{2}$$

WHERE $\hat{R}$ IS THE BEST ESTIMATE OF SYSTEM RELIABILITY

$$R_L = 1 - \hat{Q} - t_\beta \sqrt{\prod_i^n \left(\left[\hat{q}_i + (N_i - 1)\, q_i^2\right]/N_i\right) - \hat{Q}^2} \tag{3}$$

WHERE R_L IS THE β% LOWER BOUND, DETERMINED BY t_β,
AND $\hat{Q} = \prod_i^n \hat{q}_i$

Fig. 11 - Determination of the β% lower bound on system reliability (parallel arrangement).

$$\beta = 1 - \int_{t_\beta}^{\infty} \frac{1}{\sqrt{2\pi}}\, e^{-t^2/2}\, dt$$

β	t_β	β	t_β
.99	2.326	.80	0.842
.95	1.645	.70	0.542
.90	1.281	.60	0.253

Fig. 12 - Normal probability integral.

$$\text{I.}\quad dm \ln \frac{1-P_1}{1-P_2} + gm \ln \frac{P_1}{P_2} = \ln \frac{1-\alpha}{\beta}$$

$$\text{II.}\quad dm \ln \frac{1-P_1}{1-P_2} + gm \ln \frac{P_1}{P_2} = \ln \frac{\alpha}{1-\beta}$$

WHERE P_1 - MINIMUM ACCEPTABLE RELIABILITY,
P_2 - RELIABILITY DESIGN OBJECTIVE,
α - CONSUMERS RISK
β - PRODUCERS RISK,
dm - DEFECTIVE UNITS IN FIRST "m" TESTED,
AND gm - GOOD UNITS IN FIRST "m" TESTED.

Fig. 13 - Sequential analysis, equations.

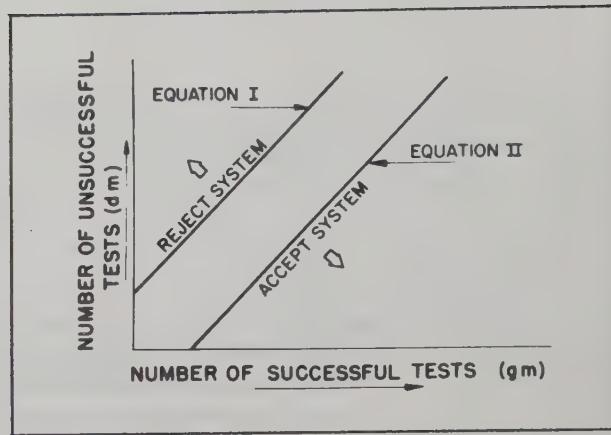

Fig. 14 - Sequential analysis, graphical representation.

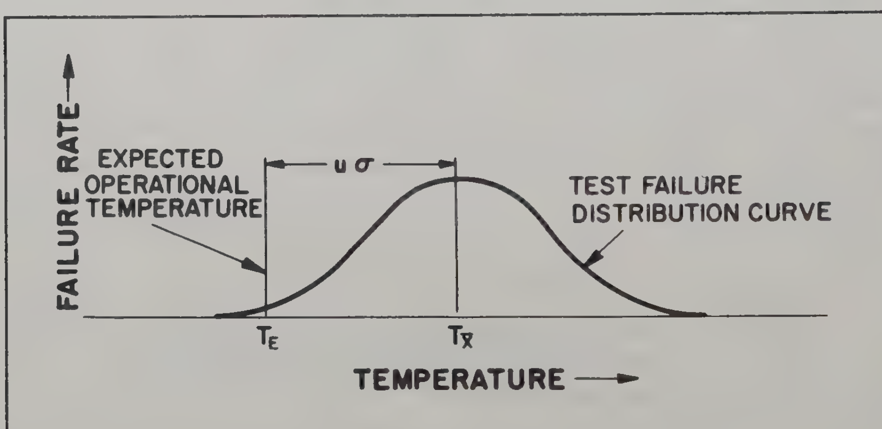

Fig. 15 - Failure distribution curve.

correspond to some point on the graph. If neither bound is crossed in the sampling process, the testing is continued. When a bound is crossed, testing is discontinued and the equipment accepted or rejected as a result of the bound crossed. This sequential sampling procedure is usually desirable where high equipment reliability is required but where time and cost considerations limit demonstration of the required reliability.

Sampling by Variables

Where available sample size does not allow for an adequate attribute testing plan, other techniques to determine equipment reliability must be used. One of the most popular of these sampling techniques is testing to failure. This method provides confidence values which are not usually obtainable under tests to specified limits. In testing to failure, one of two basic patterns is followed: using time and stress as parameters, either one can be varied with the other held constant. Figure 15 illustrates the testing to failure methodology. In this example the equipment failure pattern is determined as a function of environment. Utilizing this failure distribution curve, the probability of equipment failure at the expected operational environment can be determined.

In determining component reliability, various parametric and nonparametric sampling techniques can be employed depending on the equipment to be tested. Often in a system development program, component reliability must be estimated before reliability testing is possible. In predicting this component reliability, there are several information sources available. These include performance test data, existing operational data and component design data which are detailed in Fig. 16.

System reliability can be determined through systems level testing, component level testing as previously discussed or a combination of systems level and component level testing (Fig. 17). Testing the complete system, where possible, is desirable because system level testing provides data on component interactions, a missing element in component testing. However, system reliability can be predicted by utilizing component reliability data together with an adequate analysis of component and subsystem reliability relationships.

A. TEST DATA

1. OBJECTIVES OF TEST
2. TEST CONDITIONS
3. TEST RESULTS

B. EXISTING OPERATIONAL DATA

1. PREVIOUS OPERATIONAL USAGE
2. ENVIRONMENT IN WHICH USED
3. PERFORMANCE RECORD

C. COMPONENT DESIGN DATA

1. DESIGN LIMITS AND TOLERANCES
2. OPERATIONAL REQUIREMENTS
3. ENVIRONMENTAL REQUIREMENTS

Fig. 16 - Data requirements.

I SYSTEMS LEVEL TESTING

II COMPONENT LEVEL TESTING

A. ESTABLISH COMPONENT RELIABILITY

B. UTILIZE COMPONENT RELIABILITY DATA TO PREDICT SYSTEM RELIABILITY

1. DETERMINE COMPONENT AND SUB-SYSTEM RELIABILITY RELATIONSHIPS

2. ESTABLISH ANALYTICAL METHODS FOR PREDICTING SYSTEM RELIABILITY FROM COMPONENT DATA.

III COMBINATION OF SYSTEMS LEVEL AND COMPONENT LEVEL TESTING.

Fig. 17 - Establishing system reliability.

CONCLUSION

The reliability problem in system development can only be solved with a well designed reliability program. This program must provide for accrual of reliability knowledge, reliability indoctrination and reliability improvement. The latter can only be obtained through adequate quantitative analysis and evaluation of available component and system data. Although many methods of determining equipment reliability are available, the optimum method must be tailored for the specific system and development program.

A SEQUENTIAL TEST FOR COMPARING COMPONENT RELIABILITIES

C. F. Stevens
Applied Electronics Laboratories
The General Electric Company Limited
Stanmore, England

Summary -- A statistical technique is presented by means of which the reliabilities of two types of component (or equipment) may be compared, and a decision made in favor of one type or the other. The technique differs from those previously devised for this purpose in that the assumptions on which it is based are less restrictive and thus more likely to be satisfied in practice.

INTRODUCTION

The purpose of the tests described in this report is to compare the reliabilities of two different types of component, the outcome being a decision in favor of one type or the other. For example, a component manufacturer might wish to compare a redesigned component with the standard version to see whether a significant improvement had been achieved, or an equipment manufacturer might wish to choose between two types of component for use in his equipment. Possible applications of the test, however, are not confined to engineering. It could be used, for example, to test whether people vaccinated against a disease had a significantly smaller chance of contracting the disease than unvaccinated people.

In component applications, the experimental setup envisaged is of the two groups of components, not necessarily of equal size, being exposed simultaneously to the same experience (e.g., heat and vibration). A test to the same end has been devised by Epstein[2] but the assumption of exponentially distributed failure ages on which it is based is more restrictive than the assumption necessary in the present case.

The test is sequential in nature, which means that the results are assessed as they arise, permitting the experiment to stop as soon as enough data is available to reach a decision. Wald[5] has shown that sequential probability-ratio tests, of which this is one, are more efficient for deciding between specified alternatives than any other type of test. The point is not stressed, however, since the author is unable to find any simple nonsequential test which fulfills the same functions.

COMPARISON OF RELIABILITY

The Hazard Rate

The measure of reliability on which comparisons between component types are based in this report is the "hazard rate" $Z(x)$, sometimes called the instantaneous conditional failure rate or conditional density function. This is defined as

$$Z(x) = \lim_{\delta x \to 0} \left\{ \frac{\text{Probability that a component functioning at time } x \text{ will fail in the interval } (x, x+\delta x)}{\delta x} \right\}. \tag{1}$$

The concept of hazard rate, which is equivalent to the actuaries "force of mortality" is discussed in papers by Davis[1] and Kao[3].

A widely used model for the analysis of failure data is that in which the hazard rate is assumed to be a constant; i.e., $Z(x) = \lambda$. This is equivalent to assuming that the age of a component does not affect its liability to failure, at least within certain age limits. This assumption has been studied in some detail by Davis, by the analysis of actual failure data from a variety of sources. He concluded that in general the hypothesis of constant hazard rate was adequate to describe the failure of complex assemblies operating under a diversity of conditions, but that it was not adequate for some types of components of a more uniform nature or operating under more uniform conditions.

The failure of electronic valves on life tests has been similarly studied by Kao, who concluded that their behavior could best be described by a hazard rate of the form

$$Z(x) = (1.7)\, x^{0.7} c \tag{2}$$

in which c is some constant, characteristic of the population concerned.

In many cases one suspects that the popularity of the assumption of constant hazard rate has two causes -- the relative ease with which its mathematical consequences may be derived, and the difficulty of obtaining a "significant" departure from this hypothesis with a limited sample of data when the departure is not gross. While not wishing to decry the hypothesis of constant failure rate, therefore, it is suggested that it may be sounder in practice to use statistical techniques not dependent on this hypothesis where such can be found. This is possible when comparison between two component types is desired.

Comparison of Hazard Rates

Before proceeding to indicate the nature of the assumption on which the present method is based, the terms "experimental type" and "standard type" should be introduced to distinguish between the two component types. These terms are, however, only convenient labels, and there is no reason why the test should not be used to compare two types which are both experimental or both standard. Parameters and functions associated with each type will be distinguished by the suffixes E and S. In this notation, the two hazard rates are $Z_E(x)$ and $Z_S(x)$.

No assumption is made about the precise functional form of $Z_E(x)$ and $Z_S(x)$; it is merely assumed that they are at all times in some constant ratio θ. That is

$$\frac{Z_E(x)}{Z_S(x)} = \theta \tag{3}$$

where θ is independent of x. This condition is obviously satisfied when both hazard rates are constant, for then their ratio is also constant. It will also be satisfied when the hazard rates are given by Eq. (2), for then, if the two types are characterised by constants c_E and c_S,

$$\frac{Z_E(x)}{Z_S(x)} = \frac{(1.7)\, x^{0.7}}{c_E} \Big/ \frac{(1.7)\, x^{0.7}}{c_S} = \frac{c_S}{c_E} \tag{4}$$

which is independent of x.

An interesting example of some actual data which satisfies the condition of Eq. (3), at least approximately, is the mortality data for males and females shown in Fig. 1, which was obtained from the General Register Office.[4] By "rate of mortality" is meant the probability that a person of exact age x (years) will die before attaining age x + 1 (years), and is therefore a close approximation to the hazard rate defined by Eq. (1). Figure 1 shows that over a wide age-range the rate for males is 20 per cent to 40 per cent above the female rate. The data are plotted on a logarithmic scale in order to bring out this constant of proportionality. The adoption of the assumption of a constant ratio is advocated because it is obviously more general than any assumption of a particular functional form of hazard rate, and because it is satisfied by at least two functional forms which have been suggested for the description of observed data.

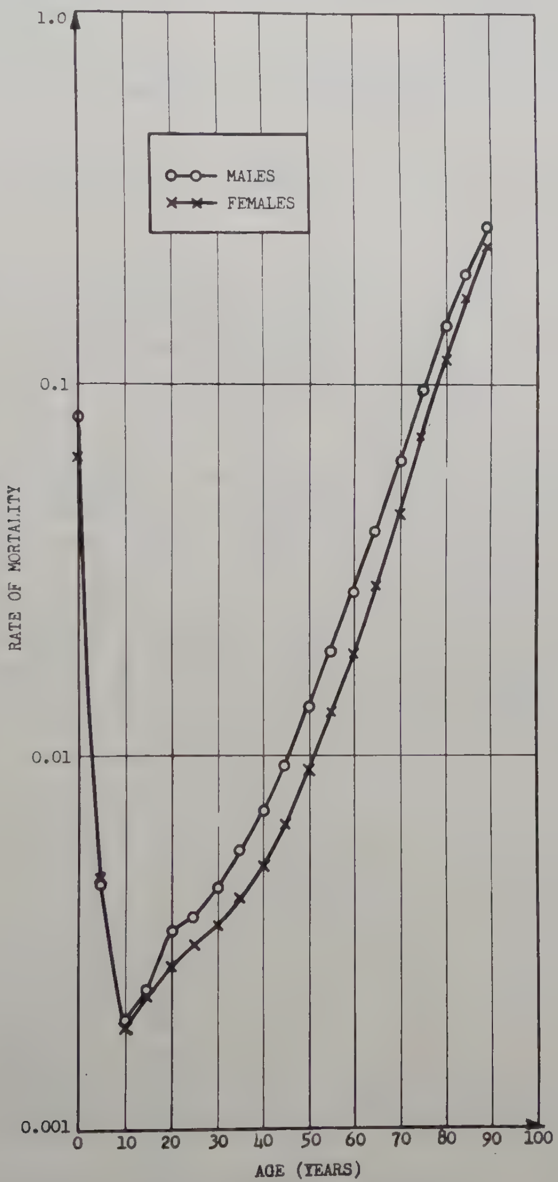

Fig. 1 - Rates of mortality, males and females, England and Wales 1921.

SEQUENTIAL TEST PROCEDURE

The ratio $\theta = Z_E(x)/Z_S(x)$ is by definition the ratio of the hazard rates of the two types, and therefore a direct comparative measure between the two types. The alternative hypotheses to be tested are therefore framed in terms of θ.

To specify a sequential test, four quantities -- θ_0, θ_1, α and β -- must be chosen beforehand. Their definitions are:

θ_0 = a number such that for values $\theta \geq \theta_0$, the <u>standard</u> type of component would be preferred;

θ_1 = a number such that for values $\theta \leq \theta_1$, the <u>experimental</u> type of component would be preferred;

α = the maximum risk of erroneously preferring the experimental type when the standard is in fact superior;

β = the maximum risk of erroneously preferring the standard type when the experimental is in fact superior.

The basis of the sequential test is the sequence of failures observed in the course of the experiment. A specimen sequence is S S E S E S ..., meaning that the first and second failures were of the standard type, the third of the experimental type, and so on. For the present purpose, no information is required on the times at which the failures occur.

Other quantities of importance in the sequential procedure are the numbers of survivors (or, at risk) at each stage of the experiment. We accordingly define $n_{E,j}$ and $n_{S,j}$ as the numbers of each type at risk before the occurrence of the j th failure. In this notation $n_{E,1}$ and $n_{S,1}$ are the numbers of each type at the commencement of the life test. In practice the numbers $n_{E,j}$ and $n_{S,j}$ are constructed successively from the initial values. Suppose a life test were started with twenty components of each type, and the failure sequence were S S E S E S ..., the procedure for calculating values of $n_{E,j}$ and $n_{S,j}$ would be that shown in Table I.

TABLE I

Failure No. (j)	Failure Type	$n_{E,j}$	$n_{S,j}$
1	S	20	20
2	S	20	19
3	E	20	18
4	S	19	18
5	E	19	17
6	S	18	17
.	.	.	16
.	.	.	.
.	.	.	.

The experiment should normally be started with equal numbers of each type. However, this is not a necessary condition, and if for any reason it is either necessary or desirable to start with unequal numbers, the computational procedure described below can still be carried out. Let X_j be used to denote symbolically the nature of the j th failure; i.e., $X_j = E$ or $X_j = S$. We require the computation of quantities $\lambda_j(X_j)$, which are defined as

$$\lambda_j(S) = \frac{1 + \theta_o n_{E,j}/n_{S,j}}{1 + \theta_1 n_{E,j}/n_{S,j}} \text{ (i.e., if } X_j = S) \tag{5}$$

$$\lambda_j(E) = \frac{\theta_1}{\theta_o} \lambda_j(S) \text{ (i.e., if } X_j = E). \tag{6}$$

As successive failures occur, the quantities $\lambda_1 (X_1), \lambda_2 (X_2), \lambda_3 (X_3),\ldots$ etc. are computed and their running product L_j formed. After the j th failure has occurred,

$$L_j = \lambda_1(X_1), \lambda_2(X_2)\ldots\lambda_j(X_j) \tag{7}$$

or,

$$\left.\begin{aligned} L_j &= \lambda_j(X_j)L_{j-1} \quad (j \geq 1) \\ L_o &= 1 \end{aligned}\right. \tag{8}$$

This completes the step-by-step computation necessary for the test. Before the commencement of the test, stopping points A and B should be computed,

$$A = \frac{(1 - \beta)}{\alpha} \tag{9}$$

$$B = \frac{\beta}{(1 - \alpha)}. \tag{10}$$

As each value L_j (j = 1,2,... etc.) is obtained, it should be tested in the following inequalities:

1. if $B < L_j < A$, insufficient information has been accumulated, and the test should proceed;

2. if $L_j \geq A$, the test should stop and preference expressed for the <u>experimental</u> type;

3. if $L_j \leq B$, the test should stop and preference expressed for the <u>standard</u> type.

A proof of the method described is contained in the Appendix.

Termination of the Test

There is a possibility that the numbers of one type of component may become exhausted before a decision to end the test has been reached. Two alternatives

are then available, the choice between them being a matter for the experimenter. The decision may be based on the available evidence up to this point, as discussed by Wald, who calls this "truncation." The practical effect of this is that the risks of error associated with this procedure are larger than the initially prescribed risks α and β. Or the test may be at any time started afresh with any number of each type of component, the likelihood-ratios L_j being carried on from the point reached by the first test.

The second alternative is the only way in which the initial objects of the test, as specified by θ_1, θ_o, α, β, may be achieved, and is therefore preferable on that account. Practical considerations may, however, necessitate the adoption of the first alternative.

Average Sample Number and O.C. Curves

Due to the analytical complexity of the problem, the author is unable to see any means of deriving mathematical expressions for the A.S.N. and O.C. curves. If a digital computer is available, however, it is feasible to simulate the test process a large number of times by a Monte Carlo method, thus obtaining any information required about the mode of operation of the test.

WORKED EXAMPLE

The foregoing techniques have been illustrated by means of a worked example. The data was obtained from laboratory vibration life tests on two types of electronic valve. The experimental type was designed to stand vibration conditions better than the standard type with which it was compared. In these circumstances the following parameter values were chosen:

$$\theta_o = 0.9$$
$$\theta_1 = 0.3$$
$$\alpha = 0.05$$
$$\beta = 0.10$$

The significance of the values chosen for θ_o and θ_1 is as follows. If $Z_E/Z_S = 0.9$ then the experimental type, while slightly better than the standard type, is not deemed sufficiently superior to justify its adoption in preference to the standard type. If, on the other hand, $Z_E/Z_S = 0.3$, then the experimental type is deemed markedly superior to the standard type and the decision should be made in its favor.

Twelve valves of each type were put on test, so that $n_{E,1} = n_{S,1} = 12$. The sequence of failures observed was S S S S S S E S S E. It would be tedious to describe the sequence of computations, so comment is confined to a number of special points.

1. Before the test starts the stopping points A and B should be computed.

$$A = \frac{(1-\beta)}{\alpha} = 18$$

$$B = \frac{\beta}{(1-\alpha)} = 0.1053$$

2. The numbers $n_{E,j}$ and $n_{S,j}$ are computed successively from the observed failure sequence. A useful check is that

$$n_{E,j} + n_{S,j} + j = n_{E,1} + n_{S,1} + 1 \ (j = 1,2\ldots).$$

3. In calculating $\lambda_j(X)_j$, it is recommended that $\lambda_j(S)$ be computed first, whatever X_j is. Then, if $X_j = E$, $\lambda_j(E)$ can be obtained from the relation of Eq. (6),

$$\lambda_j(E) = (\theta_1/\theta_0)\lambda_j(S).$$

4. The likelihood-ratio is then computed as the running product

$$L_j = \lambda_1(X_1)\lambda_2(X_2)\ldots\ \lambda_j(X_j).$$

5. At each stage, L_j is compared with stopping points A and B (see 1 above). If

   a. $B < L_j < A$, the test continues;

   b. $L_j \geqslant A$, the test ends with a decision in favor of the experimental type;

   c. $L_j \leqslant B$, the test ends with a decision in favor of the standard type.

The test in this example stops at the 9 th failure with a decision in favor of the experimental type. It is of interest to note that with the values of θ_0, θ_1, α, β chosen for this test, this sequence of 8/12 S-failures and 1/12 E-failures is only just sufficient to insure a decision in favor of the experimental type.

APPENDIX

Derivation of the Sequential Test Procedure

Let $Z_E(x)$ and $Z_S(x)$ be the hazard rates of the experimental and standard types, respectively. Let the numbers of each type at a given instant x be n_E and n_S. Then

$$Pr\left\{\text{An E-failure in } (x,x+\delta x)\right\} = n_E Z_E(x)\delta x + O(\delta x^2) \quad \text{(A1)}$$

$$Pr\left\{\text{An S-failure in } (x,x+\delta x)\right\} = n_S Z_S(x)\delta x + O(\delta x^2). \quad \text{(A2)}$$

Suppose the time axis to be divided up into a number of very short intervals of length δx. Only those intervals in which failures occur will be considered, and since the interval length δx may be made arbitrarily small we may confine our attention to those intervals in which a single failure occurs, the probability of more than one failure being of the order of $(\delta x)^2$. The failures will be called E-type or S-type failures, depending on the type of the failed component. We shall write p(E) and p(S) to denote the probabilities of an E-type or S-type failure, conditional on a failure of some type having occurred. From Eqs. (A1) and (A2)

$$P(E) = \frac{n_E Z_E(x)}{n_E Z_E(x) + n_S Z_S(x)} \tag{A3}$$

$$P(S) = \frac{n_S Z_S(x)}{n_E Z_E(x) + n_S Z_S(x)} \tag{A4}$$

(these being the limiting values when the interval length δx is made arbitrarily small).

Writing

$$\frac{Z_E(x)}{Z_S(x)} = \theta \ , \tag{A5}$$

we then have, finally,

$$p(E) = \frac{\theta n_E/n_S}{1 + \theta n_E/n_S} \tag{A3.1}$$

$$p(S) = \frac{1}{1 + \theta n_E/n_S} \ . \tag{A4.1}$$

Eqs. (A3.1) and (A4.1) show the dependence on θ of the conditional probabilities of E-type or S-type failures, given that failure of some type has occurred; to emphasise this dependence we shall when necessary, write $p(E|\theta)$ and $p(S|\theta)$.

Let X_j denote the nature of the j th failure (symbolically, either $X_j = E$ or $X_j = S$). The result of a test may be expressed as an observed sequence X_1, X_2,...etc.

We are concerned with testing the hypothesis $\theta = \theta_o$ against the alternative hypothesis $\theta = \theta_1$. The likelihood-ratio L_j of a set of observations X_1, X_2,... X_j is

$$L_j = \frac{p(X_1, \ldots X_j | \theta_1)}{p(X_1, \ldots X_j | \theta_o)} \tag{A6}$$

$$= \prod_{t=1}^{j} \frac{p(X_t | \theta_1)}{p(X_t | \theta_o)} \tag{A7}$$

Write

$$A = \frac{(1 - \beta)}{\alpha} \tag{A8}$$

$$B = \frac{\beta}{(1 - \alpha)} \ . \tag{A9}$$

Wald[5] has shown that the sequential probability-ratio test with the prescribed risks of error may be conducted by computing successive values of L_j ($j = 1, 2, \ldots$etc.) and using the following rules:

1. if $B < L_j < A$, insufficient information is available to make a decision and the test should proceed;

2. when $L_j \geq A$ the test should stop and the decision made in favor of $\theta = \theta_1$;

3. when $L_j \leq B$ the test should stop and the decision made in favor of the null-hypothesis, $\theta = \theta_o$.

The values of the constants θ_o and θ_1 should be chosen so that $\theta = \theta_o$ implies a preference for the standard type and $\theta = \theta_1$ implies a preference for the experimental type.

The sequential test procedure therefore requires a step-by-step evaluation of L_j, to which we now turn.

Let

$$\lambda_j(X_j) = \frac{p(X_j \mid \theta_1)}{p(X_j \mid \theta_o)} . \qquad \text{(A10)}$$

Then, from Eq. (A7),

$$L_j = \prod_{t=1}^{j} \lambda_t(X_t) . \qquad \text{(A11)}$$

Let $n_{E,j}$ and $n_{S,j}$ be the numbers of each type of component functioning prior to the occurrence of the j th failure.

From Eqs. (A3.1) and (A10) we find

$$\lambda_j(E) = \frac{p(X_j = E \mid \theta_1)}{p(X_j = E \mid \theta_o)}$$

$$= \frac{\theta_1}{\theta_o} \cdot \frac{1 + \theta_o n_{E,j}/n_{S,j}}{1 + \theta_1 n_{E,j}/n_{S,j}} . \qquad \text{(A12)}$$

Similarly

$$\lambda_j(S) = \frac{1 + \theta_o n_{E,j}/n_{S,j}}{1 + \theta_1 n_{E,j}/n_{S,j}} . \qquad \text{(A13)}$$

If the j th failure is of type E, we compute $\lambda_j(E)$; if of type S, we compute $\lambda_j(S)$. The current value of the likelihood-ratio L_j may then be calculated from Eq. (A11), which may be written

$$L_j = \lambda_j(X_j) \cdot L_{j-1}. \qquad \text{(A14)}$$

L_j is therefore built up as a running product. The test terminates when either $L_j \geq A$ or $L_j \leq B$ for the first time, as described.

TABLE II

Worked Example

$\theta_o = 0.9 \quad \alpha = 0.05 \quad A = 18$

$\theta_1 = 0.3 \quad \beta = 0.10 \quad B = 0.1053$

Failure No. (j)	Failure Type (X_j)	$n_{E,j}$	$n_{S,j}$	$n_{E,j}/n_{S,j}$	$1+\theta_o\frac{n_{E,j}}{n_{S,j}}$	$1+\theta_1\frac{n_{E,j}}{n_{S,j}}$	$\lambda_j(S)$	$\lambda_j(E)$	$\lambda_j(X_j)$	L_j
1	S	12	12	1.000 000	1.900 000	1.300 000	1.461 538	-	1.4615	1.4615
2	S	12	11	1.090 909	1.981 818	1.327 273	1.493 150	-	1.4932	2.1823
3	S	12	10	1.200 000	2.080 000	1.360 000	1.529 412	-	1.5294	3.3376
4	S	12	9	1.333 333	2.199 999	1.399 999	1.571 429	-	1.5714	5.2447
5	S	12	8	1.500 000	2.350 000	1.450 000	1.620 690	-	1.6207	8.5001
6	S	12	7	1.714 286	2.542 857	1.514 286	1.679 245	-	1.6792	14.2734
7	E	12	6	2.000 000	2.800 000	1.600 000	1.750 000	0.583 333	0.5833	8.3257
8	S	11	6	1.833 333	2.649 999	1.549 999	1.709 678	-	1.7097	14.2344
9	S	11	5	2.200 000	2.980 000	1.660 000	1.795 181	-	1.7952	25.5536
10	E			Test stopped after 9th failure, since			$L_9 \geq A = 18$			

Notes:

$$\lambda_j(S) = \frac{1+\theta_o\frac{n_{E,j}}{n_{S,j}}}{1+\theta_1\frac{n_{E,j}}{n_{S,j}}}$$

$$\left.\begin{aligned}\lambda_j(X_j) &= \lambda_j(S) \text{ if } X_j = S \\ &= \lambda_j(E) \text{ if } X_j = E\end{aligned}\right\}$$

$$\lambda_j(E) = (\theta_1/\theta_o)\lambda_j(S)$$

$$L_j = \lambda_1(X_1)\lambda_2(X_2) \ldots {}_j(X_j)$$

BIBLIOGRAPHY

1. Davis, D. J.: "An analysis of some failure data." Journal of the American Statistical Association, 47:113-150 (June 1952).

2. Epstein, B.: "A sequential two sample life test." Journal of the Franklin Institute, 260:25-29 (1955).

3. Kao, J. H.: "A new life-quality measure for electron tubes." Trans. IRE, PGRQC-7, PGRQC-7, pp. 1-11 (April 1956).

4. General Register Office: The Registrar General's Decennial Supplement, England and Wales, 1921, Part I, Life Tables. Her Majesty's Stationery Office, London, 1927.

5. Wald, A.: "Sequential Analysis." New York, John Wiley and Sons, 1947.

6. Weibull, W.: "A statistical distribution of wide applicability." Journal of Applied Mechanics, 18:293-297 (Sept. 1951).